Good things come to those who weight: evidence integration and decision termination in human choices

Hannah Tickle

Department of Experimental Psychology

University College London

Thesis submitted for the degree of

Doctor of Philosophy

June 2018

Declaration

I, Hannah Pamela Tickle, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

HANNAH TICKLE

Acknowledgements

There are numerous people who I need to thank for providing practical and/or moral support during the writing of this thesis (a.k.a. “pile of bollocking rubbish about brains!”, as Aunty Pen put it three ports down after Christmas dinner, and the name stuck). I realised when thinking about who to acknowledge that it is impossible to mention everyone personally, which I guess is a sign of how well supported I’ve felt. All was appreciated.

I have to start with Nanny, who (with Mum’s help) instilled the tea addiction that probably itself deserves an acknowledgement. Who, in being prepared to drop the c bomb on the scrabble board just to get the triple word score, taught me the importance of courage and determination to succeed. Thank you for your all-encompassing, overwhelming love and kindness. I remember you with love every day.

And to the rest of the family - to Mum, for nurturing the sense of Girl Power, to Dad, for sage life advice, to Lizzie, for acting as my common sense, to Joseph, for helping me to see the world from a different perspective. To Richard, for being a motivating presence during last minute thesis stress (finally time for the champagne). Mum pointed out the other day how unusual it is to have such a tight knit extended family, so I can’t possibly name everyone else individually, but above all I want to acknowledge all the family (and friends who were close enough to be family) who stayed with me through those years during which even achieving A levels seemed unrealistic, who stuck around at such cost with such little gain for so long. All I’ve achieved has only been possible because of you.

To my excellent friends – Helen, for patiently wondering when I’d get my act together with this thesis, and for not judging when a mere three years later I finally figured out and exclaimed to her with wonder that it’s actually quite time consuming. To Sam, for believing that it was about the swans (and it is!! They are in here somewhere). To Alex, for being there for me, and indeed all women. To Tilly, for being my PhD buddy, to Hannah, for being my science buddy (missing you), to Chloe for giving me (and many in the lab) some Shakespeare to enrich ourselves. Thank you all for being wonderful friends and putting up with the perennial student in your midst (and thanks Tilly for providing back-up in that respect).

To everyone in the Summerfield Lab, past and present – can’t mention you all personally but you are all wholeheartedly appreciated. Janto and Konstantinos, thank you for your contributions to the work here. Fabrice thank you for being the definition of a good egg, and Keno for helping me coerce others into pub trips. Lizzie M, thank you for helping me persevere in the face of adversity and for being an invaluable source of scientific and feminist support. Jan, thank you for continual thoughtfulness and unswerving helpfulness. I don't think I've ever asked you something about science that you've not gone out your way to answer, and it is tremendously appreciated.

At UCL (and miscellaneous science friends) – Tom (and Steve) – starting to share the office (and cups of tea and endless chats and bake off) with you was the first time I actually started to enjoy coming into Bedford Way. Sabine, my PhD sister, I thoroughly enjoyed our excellent and increasingly manic months of office companionship. To all others in that room, you made the PhD what it was. To Poppy for countless extremely funny experiences, for reams of chat-based PhD support, for possessing the

key to that lifesaving dungeon sofa room after too much post-conference revelry (sorry for breaking your bike).

And finally, of course, to Maarten and Chris. Maarten, thank you for taking a chance on me when I stumbled into your office in search of a supervisor. Thanks for patiently accepting my totally wild and inaccurate speculations on when I may finish these chapters/the thesis in general, for explaining stuff that I tell you I understand afterwards, and then forget to implement it and we start all over again. For teaching the most useful stats course I've ever encountered. Thanks for being laid back, but always being there to answer questions. Everything you've done is appreciated.

And Chris – you always accused me of looking bored in your tutorials, but hopefully the fact I'm still working with you 9 years later is sufficient proof that I wasn't. On day 1 of my PhD I hadn't found Maarten yet and was pretty disillusioned, but you and your lab welcomed me and I've never stopped being grateful. I'm grateful for everything you've done (apart from banning Harry Potter references in the thesis, but I guess you can't prevent this one): for the near death experiences (never letting you drive a minibus again), for consistently predicting an opposite pattern of results to what we find, for playing devil's advocate to any opinion I've ever raised in a pub, for the thousands of hours you've dedicated to helping me mature as a scientist, and for being a hugely valued supervisor and friend.

Long abstract

Perceptual decision-making describes the processes by which sensory information is recognised, evaluated and combined before making a commitment to a course of action. The goal of this thesis is to understand the neural and computational mechanisms underlying human perceptual decisions. Good decisions are made when all the available evidence is taken into account, and allowed to influence choice in proportion to its reliability. The first experimental chapter describes a categorisation task employed to investigate how information is integrated and employed according to its reliability during sequential sampling. It is observed that humans weight information approximately optimally. A subsequent experiment involving electroencephalographic (EEG) recordings elucidates a neurobiologically plausible mechanism that could give rise to this effect.

However, reliability-based evidence integration may only be possible in relatively simple decisions, when task demands are lower. Previous work investigating more challenging decisions has shown that when two alternatives are viewed in series, locally preferred alternatives are processed with higher gain ("selective integration"). Experiment 2 asks (at both the behavioural and neural level) whether this selective integration happens at the level of *attributes* - i.e. category A versus B - or *features* - i.e. sub-dimensions of each of the attributes. Finding that it occurs at the level of features, we discuss the optimality of this strategy. We show, interestingly, that whilst selective integration at the feature level is not harmful to performance, only attribute-level selectivity is actively beneficial in this context.

In everyday settings, the choice to stop integrating evidence and commit is often determined by the agent, rather than an external deadline. Experiment 3 uses a self-paced categorisation task to investigate what factors predict *when* decisions are made. The results show that decisions and their *latencies* are described by a quasi-optimal model, that times commitment in a way that depends on the evidence consistency. We show that an approximation based on normalisation can account for these findings at the computational level. This model predicts neural signals observed in humans.

Impact statement

Decisions are made collectively, individually, with wide- or narrow-reaching scope, successfully or unsuccessfully, in every walk of life at every moment of every day. The work presented in this thesis asks a fundamental question: when making decisions on the basis of the available information, which pieces of evidence do people attend to, and why do they have that focus? In the research presented, I show that individuals attend to only some of the evidence available in many situations, and specifically often focus on evidence that is in line with their current beliefs. Importantly, I outline the *mechanisms* that give rise to this behaviour: it is only through understanding the processes underpinning evidence integration that we can help people make the best choices. For example, when a nation goes to vote on some historically important issue about which it is generally ill-informed, knowing about these tendencies towards incomplete and biased evidence perusal may allow us to present evidence in a way that best mitigates against these effects, allowing individuals to make the best-informed decision possible.

In the world of mental health, disorders such as obsessive-compulsive disorder, depression and schizophrenia are characterised by atypical decision-making. Through understanding how decision-making occurs in healthy individuals, as I do here, we can begin to understand what causes choices to be different in those who suffer from these illnesses. Gaining this understanding may open up potential cognitive-behaviour treatment options, that guide the individual back towards a healthy pattern of decision making without the need for psychoactive medication, which often has far reaching negative side effects.

1. General introduction

1.1. Perceptual decision making

Mentioning to a newly-met acquaintance that your research focuses on decision making is a risky business: there generally follows an exclamation of great excitement about the clear utility of your work, followed by an inevitable anti-climax when you hasten to add that the decisions in question are perceptual in nature. To many, perceptual decision making – the detection, discrimination and categorisation of sensory information, which underpins processes from determining a person's identity, to deducing whether you have been pocket dialled or the person on the other end of the phone just has bad signal, to recognising the bus in the distance as the one you are waiting for – is just not something that they would classify as 'decision making': is it really fair, they ask, to call a process that generally occurs with such obvious ease, with little to no conscious deliberation, a "decision"?

A decision occurs every time we are compelled to choose one option from a number of alternatives. Thus objectively, yes, such perceptual processes certainly are decisions, and furthermore they are far from trivial. When I correctly recognise movement I catch in the corner of my eye as being caused by the cat walking past the doorway, the process seems simple, and yet within a few milliseconds my brain has detected motion in my peripheral vision and, based on the nature of this fleeting and ambiguous sensory signal, has ruled out other plausible explanations for its cause (e.g. reflections caused by my wristwatch catching the sunlight, another person walking by, an illusion caused by the onset of a migraine). What's more, I would argue that a process that occurs in such a streamlined manner that it doesn't even appear to penetrate our consciousness in fact makes it all the more remarkable. The apparent structure of the external world seems so obvious to us that we may forget that every waking moment, we are being bombarded with sensory information and that somehow – somehow – from this jumbled mess we are able to segment objects, notice coherence, differences, coincidences, patterns and much more in order to infer causes to these sensations. This is, I would argue, a breath-taking ability - especially when we consider that the subjective perception that we take into account everything that is 'out there' cannot possibly be true. Thus without even fully

processing the information, we pick out the relevant evidence, combine it with other pieces of information and use this to guide our decisions, often under intense time pressure.

In this thesis I will be trying to understand some of the processes that give rise to this ability. I will begin with an analogy that covers the key themes that will follow: imagine playing a game of poker at a casino. The player to your right, who has been betting with dreadful hands all evening, throws some money into the pot. You are aware that there are better hands than your own that could be out there, but given this player's tendency to bluff you match his bet. However, after the next card is revealed, a different player on your left – who is evidently a skilled poker player – makes another, similarly sized bet, and this now gives you pause. You note the cards on the table that are higher than your own, and reflect on the probability of your own hand winning. Concluding after a few moments of deliberation that she likely has a better hand than you, you fold your cards before the other players start to complain that your thinking time is holding up the game. This process, lasting only a few seconds, is an apt example of several of the phenomena that will be examined in detail in this thesis.

Firstly, to make good decisions, information must be considered in proportion to its reliability. Giving much greater credence to the raise from the skilled player than the bet from the bluffer illustrates behaviour in accordance with this principle: although the 'evidence' (the size of the bet) was equally strong in both cases, they lead to very different behaviour on your part (calling versus folding). The question of whether and how people make decisions that take into account the reliability of the information is of central interest here.

Secondly, we noted already that it is impossible to take into account all of the information that reaches our senses, but even information which is processed in conscious awareness may be ignored. When the second player made her bet, although there were several cards on the table, those that had the strongest influence on your decision to fold were the higher cards. This is despite the fact that it's entirely plausible that lower cards formed part of her winning hand. Thus a second central question I will ask is what makes us integrate some information preferentially over other pieces of information, and what impact this has on choices.

Thirdly, I will consider what drives the timing of choices. In the poker example, your decision to fold was made in a timely manner, despite there being no strict deadline imposed on choices (other than

etiquette). If a decision is made too quickly, it is likely to be bad, as there will not have been time to consider a sufficient amount of evidence. However if a decision is made too slowly, it could also be bad: the opportunity for the outcome to be beneficial may already have passed. How we strike a balance to time decisions appropriately will be a third focus here.

Finally, an overarching theme of this work is: do these processes lead to good choices? Humans are after all often criticised for making poor choices (Brexit, for example), and selective integration of only some of the information in particular seems like a potentially myopic strategy. However, we will see that defining "optimality" in the first place is a far from simple process.

In the remainder of this introduction, thus, I will start by outlining traditional definitions of "optimality", and providing evidence that people do or do not make optimal perceptual decisions. Finding conflicting evidence, I will question these traditional definitions, and suggest an alternative definition of 'optimality' that provides a unifying account of otherwise contradictory findings. According to this updated definition, an optimal decision is one that leads to the greatest chance of making an accurate choice when cognitive demand is high. Finally, I will discuss theories that account for *when* choices are made.

1.2. Do humans make good decisions?

Agents make good decisions by giving greater credence to more reliable information

Good decisions use all available evidence to inform the choice. However, the best decisions not only take into account all the available information but they also – crucially – scale the impact of evidence on choice according to its *reliability*, over and above its strength. The oft used example to illustrate what it means to separate these two factors, analogous to the poker example above, is witness testimony in a court of law: two witnesses may both give internally consistent accounts of the suspect's movements on the evening the crime was committed, thus providing equally strong evidence to the case. However, said witnesses may differ substantially in their reliability: if one is an

unconnected bystander, and the other a suspected co-conspirator of the accused, the testimony of the former should be more impactful than that of the latter. 'Reliability-weighting' or 'reliability-based integration' refers to the process by which evidence is considered in proportion to its reliability, with more reliable evidence holding greater credence.

Early work detailed the phenomenon of 'visual capture': participants judged the size of an object by viewing it through a distorting lens at the same time as feeling it (Rock & Victor, 1964). The authors showed that participants were generally unaware of the discrepancy between the two sources of information, and that judgements were overwhelmingly made on the basis of visual rather than haptic information. In terms of reliability-weighting, this reliance on vision is appropriate: our highly refined visual systems often make it the most reliable source of sensory evidence (Welch & Warren, 1986).

However, later work based on maximum likelihood estimation revealed that visual evidence only dominates when it conveys the most reliable evidence: when visual sources are corrupted and therefore unreliable, their impact is attenuated by that of other senses. Maximum likelihood estimation is a process that specifies how information from different sources (for example, sensory modalities) should be combined according to the reliability of each source. Thus it can be used to derive the 'optimal' weight a reliability-weighting observer should give each source of information in a given scenario. There is a wealth of evidence, particularly from sensorimotor tasks, to support the idea that humans behave in accordance with this process. In one pivotal study, Ernst and Banks asked participants to judge the height of a bar based on either haptic evidence, or on visual evidence in the form of a random dot stereogram, the reliability of which was varied by increasing the random displacement of dots (Ernst & Banks, 2002). Asking participants to perform these unimodal judgements allowed the authors to measure the sensory uncertainty in each modality, and therefore to determine via maximum likelihood estimation how participants should behave in subsequent multimodal trials, in which the height of the block was again judged, but this time using both sources of information simultaneously. In line with the maximum likelihood rule, participants reliability-weighted the visual and haptic information in the multimodal task. Similar behaviour has also been observed in tasks that combined vision and audition (Alais & Burr, 2004; Alais, Morrone, & Burr, 2006) and vision and proprioceptive cues (van Beers, Sittig, & Gon, 1999), suggesting that reliability-weighting of sensory information is a robust, modality independent phenomenon.

'Reliable' evidence need not just be defined according to the extent of external (sensory) uncertainty (noise) at its source. A second characterisation of reliability pertains to the number of pieces of information that give rise to a certain piece of evidence: a 4.5 star TripAdvisor restaurant review should be more impactful if it is based on an average of 2000 rather than 20 reviewers. This is because statistically, information from a larger sample provides more reliable information than that from a smaller sample, as the standard error of the mean information conveyed is inversely related to the size of the sample. In a realistic task that asked individuals to rate products from the online retail site Amazon, participants first gave their initial impression (rating out of 5) of products that are actually available from the site (e.g. a pair of headphones) based only on the picture and brief description of the product. They were then shown the same products alongside their Amazon ratings (i.e. the number of people who gave the product 1, 2, 3, 4 or 5 stars) and provided a second rating. The researchers found that participants' update of their initial impression was largest when the number of reviews that formed the Amazon rating was higher and therefore more reliable, which is what would be expected if observers were indeed reliability-weighting the information (De Martino, Bobadilla-Suarez, Nouguchi, Sharot, & Love, 2017).

The example of a court of law above highlights a third type of reliability: social trustworthiness. In macaques, a monkey's position in the social hierarchy determines their relative priority in accessing food (amongst other things). As such, information about higher (rather than lower) status animals should be integrated preferentially as it is a more reliable source of pertinent information, which is exactly what research has shown: macaques forego rewards to view faces of higher status monkeys but require payment to view faces of lower status monkeys (Deaner, Khera, & Platt, 2005; Rudebeck, Buckley, Walton, & Rushworth, 2006). In human research, individuals acquire and employ information about the social reliability of a confederate who advises them during a perceptual classification task via an associative learning mechanism similar to that believed to underlie reward based learning (Behrens, Hunt, Woolrich, & Rushworth, 2008). This latter experiment is particularly important as it gives an insight into the processes used to derive the weight given to information in the first place, mirroring findings from the developmental literature that indicate that this information is learned over time (Petrini, Remark, Smith, & Nardini, 2014).

Whilst differing slightly in the specifics of their definitions, a key unifying feature of the research described is the dissociable impact of evidence reliability and strength on choice: for example in Ernst

and Bank's task, both the visual and haptic cues may have suggested that one bar was (e.g.) 2cm taller than the other, but the signal reliability is what caused information from one modality to carry greater weight. In the Amazon task, two reviews that both averaged 4 stars would have been indistinguishable except for their reliability. If reliability-weighting describes human behaviour well, we may therefore reasonably expect dissociable neural correlates of evidence strength and reliability during similar tasks. This phenomenon has indeed been observed in several studies. When participants were asked to make judgements about the average colour of an array of shapes, decreasing evidence strength (the average colour of the shapes) and increasing evidence variability (the width of the distribution from which the colour for each shape in the array was drawn, see **Figure 1.1**) not only independently negatively influenced accuracy and response latencies, but also predicted opposite response profiles in the medial prefrontal cortex (mPFC) such that activity scaled positively with the mean and negatively with the variance of the information (Michael, de Gardelle, Nevado-Holgado, & Summerfield, 2015). The neural dissociability of these two factors was similarly evident using the same task with concurrent electroencephalogram (EEG) recordings (Michael, 2015).

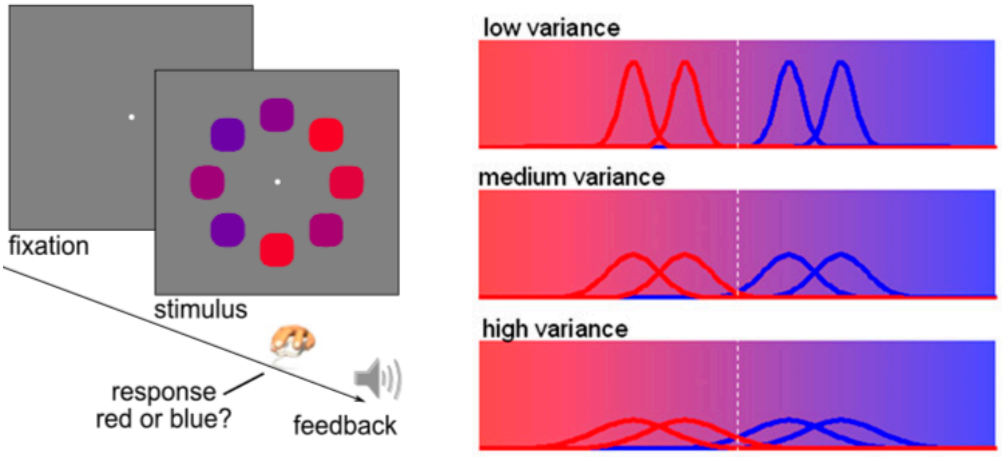

Figure 1.1: Illustration of typical stimuli used in perceptual averaging tasks, and the underlying stimulus distributions. In the task illustrated (left panel), participants were required to report the average colour of the eight shapes. As the right panel shows, when the mean of the distribution was closer to the boundary (white line), the choice was harder because the average colour information (depicted in the background of the figure) was ambiguous. When the distribution variance was wider, difficulty was increased by the range of colours that could feature in a given sample. Thus the level of mean and the level of variance had orthogonal influences on task difficulty. Figure taken from de Gardelle and Summerfield (2011).

Two important themes have emerged over the course of the discussion so far: people and monkeys reliability-weight information in order to make the best choices, and the impact of evidence reliability is dissociable from the impact of evidence strength both behaviourally and neurally. As a whole, this body of evidence supports the idea that humans (and non-human primates) do indeed make 'good' decisions. However, this account of what constitutes a 'good' decision arguably holds for the simplest decisions only, as it overlooks a key factor in the general decision making process: the influence of pre-existing knowledge. To see whether individuals optimally incorporate this information into choices, we turn next to work based on the theory of Bayesian optimality.

Agents make good decisions by behaving in line with the Bayesian solution

Bayesian decision theory holds that all beliefs about the world are assigned a probability (Green & Swets, 1966; Kording & Wolpert, 2006). Proponents of the theory note that uncertainty – which can be caused by a myriad of factors – is an inherent aspect of all decisions, and therefore the best decisions can be made by leveraging this uncertainty such that beliefs with higher probabilities carry more weight on the eventual choice. This concept is formalised by Bayes' rule:

$$p(h|x) = \frac{p(h) \cdot p(x|h)}{p(x)}$$

Where h is a hypothesis, and x is the noisy sensory observation, see **figure 1.2**. Bayes' theory is an extension of the work on reliability-weighting, and the example Kording and Wolpert (2006) use helps to illustrate how: in a game of tennis, you have a pre-existing idea of the locations on the court at which a serve is likely to land. These possible locations have a given mean (the average place that balls land) and variance (the range of places it could land), and as such can be represented as a probability distribution, which constitutes the prior $p(h)$. As your opponent takes the shot, you can then estimate where the ball will land based on the current sensory information x; given inherent uncertainty (from, for example, the visibility on the day, the speed of the movement) this estimate too is uncertain, and can therefore also be represented as a distribution with given mean and variance. The variance reflects your level of uncertainty, which itself is determined by the reliability of the sensory signal, and is expressed as a likelihood function, i.e. the probability of observing that sensory data under each hypothesis, $p(x|h)$. Crucially, if either distribution has a higher variance (i.e. lower reliability), it will

have a smaller influence on this combined likelihood estimate. The posterior (i.e. your final belief about where the ball will land) is then computed using Bayes' theorem above, and is driven more strongly by the more reliable source. It is this posterior that constitutes the Bayesian estimate of where the ball will land (Kording & Wolpert, 2006). Importantly, the probabilistic nature of decisions means that the Bayesian course of action will generally, but not always, lead to the desired outcome: whilst a single choice may lead to failure, in the long term accuracy (and thus, in many cases, reward) will be maximised if decisions are made in line with Bayes' theorem.

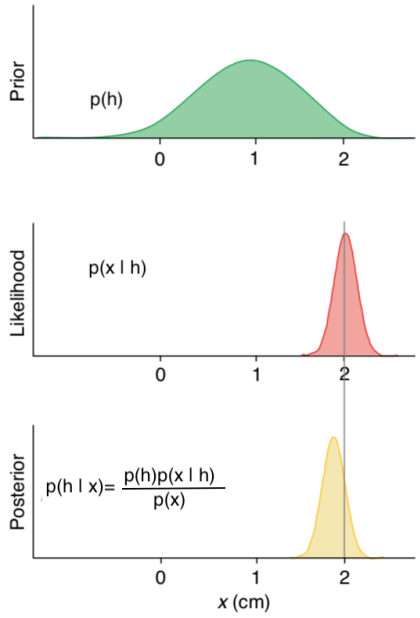

Figure 1.2: Illustration of Bayesian integration. The yellow curve at the bottom reflects the posterior, generated by combining the green curve (the prior) and the red curve (the likelihood) according to Bayes' rule. Figure taken from Kording and Wolpert (2006).

Does Bayes' theorem provide a good account of human decisions? For many researchers, the answer is 'yes'. A canonical study from Kording and Wolpert was amongst the first to demonstrate that humans optimally combine prior and current information in line with Bayes' theorem. Participants performed a task in which they were simply required to point to a target in a virtual reality environment which distorted the visual information about their arm location by an average of 1cm rightwards. On some trials, participants were given information about the true arm location mid-way through the point trajectory, with varying degrees of reliability (caused by the presence and degree of visual blurriness). Aware of this average 1cm false feedback, a Bayesian participant should compensate for the distorted visual input by altering their aim 1cm to the left (i.e. making use of the

prior information) when no mid-way feedback was provided or when the mid-way feedback was unreliable, but base their decision more on the mid-way feedback when it was less corrupted by noise. This is indeed exactly what participants did (Kording & Wolpert, 2004). Such instances of Bayesian behaviour are not limited to the sensorimotor sphere. Bayes' theorem can be applied to explain a host of behaviours: our succumbing to an illusion that causes an identical stimulus to be perceived as convex when at 0°, but concave when rotated 180°, can be explained by our Bayesian reliance on the 'light comes from above' prior, which in the vast majority of decisions leads to an optimal choice (Sun & Perona, 1998). The fact that the strength of this prior is altered by experience lends further support to Bayesian explanations, as the posterior (which incorporates previous experiences) then forms the prior (Adams, Graf, & Ernst, 2004). Similarly, our altering perception of line length as a function of its orientation or context is explained by a strong prior over the regularities of structure in the world (Howe & Purves, 2002, 2005a, 2005b). Beyond the sensory sphere, decisions made in line with Bayes' theorem have also been shown in higher-order cognition: highly abstract problems that involved inferring the prior probability distribution of questions as broad as the length of a Pharaoh's reign given the amount of time he or she had reigned already, or how long a cake will take to bake given the time it has already been in the oven, lead to answers in line with a Bayesian calculation (Griffiths & Tenenbaum, 2006). Taken together, these results provide clear evidence that patterns of behaviour in a range of tasks are in line with those that would be observed if individuals were implementing Bayes' theorem, and are a small selection of the pieces of evidence that explain why the notion that humans are "Bayes optimal" has become so dominant in the field.

Marr's levels of analysis assert that information processing systems must be understood at three levels: the computational level, which asks what the system aims to do, the algorithmic level, which asks via what processes the system does this, and the implementational level, which questions via what 'hardware' (such as neurons and neural structures) this process is implemented (Marr, 1981). Bayes' theory accounts for the computational problem to be solved: decisions must be made despite inherent uncertainty. It clearly accounts for the algorithmic process via which this may occur: it occurs according to Bayes' rule. However, we are left questioning how, at the implementational level, Bayes' theorem is manifested? Indeed, it is worth considering what type of evidence could theoretically support Bayes' theorem at the implementational level. Arguably this evidence must be neural, given that disparate mechanisms can all lead to highly similar patterns of behaviour. Evidence for the neural implementation of Bayes' theorem may involve (but not be limited to): i) evidence for explicit

representation of uncertainty or probabilities, ii) evidence for dissociable influences of prior and current sensory information, alongside iii) a mechanistic account of how they merge. If we are unable to reveal conclusive evidence then arguably, the theory falls short, despite its overwhelming popularity. (See (Jones & Love, 2011)).

There are several pieces of evidence that suggest there is an explicit neural representation of uncertainty. Platt and Glimcher trained monkeys to saccade to one of two possible locations, each of which was associated with a given prior probability. Concurrent single unit recording revealed that the responses of neurons in the lateral intraparietal cortex (LIP) were proportional to the prior probability of making a saccade to the location of that particular cell's receptive field. Additionally, the probability of reward associated with each saccade was manipulated, and neuronal responses were found to be proportional to this probability (Platt & Glimcher, 1999). A second study addressed the representation of uncertainty on the level of neurotransmitters, drawing a distinction between expected uncertainty – the known unreliability of predictive cues (it may rain today so I may need an umbrella), which the authors showed was signalled by acetylcholine, and unexpected uncertainty – unsignalled context switches that lead to unexpected observations (e.g. a substantial drop in the forecast reliability) and is signalled by norepinephrine (Yu & Dayan, 2004). However, these findings clearly still leave questions unanswered. Firing at a rate proportional to the probability is one thing, but we have 100 billion neurons (Herculano-Houzel, 2009), and decisions aren't made on the basis of the activity of just one of them. A key outstanding question is: how are populations of neurons able to represent the reliability of a stimulus simultaneously with its strength, as required by Bayesian solutions? Some sophisticated attempts have been made to outline plausible possibilities: Ma et al note that neuron spike timings occur at a Poisson rate (i.e. a fixed probability of a spike occurring at any given time), and as a result the mean of the firing rate is equivalent to its variance (reliability). Thus a population of neurons in theory could signal the probability associated with a stimulus via the range of responses in the population (Ma, Beck, Latham, & Pouget, 2006; Pouget, Beck, Ma, & Latham, 2013).

More direct support for the implementation of Bayesian processes in the brain is derived when looking for support for the second and third criteria, evidence for the representation and combination of both prior and current sensory information. One well-known example of a process that accounts for this is the theory of predictive coding (Huang & Rao, 2011; Rao & Ballard, 1999), the central tenet

of which is that rather than simply signalling the information from the sensory world, neurons signal the extent to which the incoming sensory information deviates from a prediction (the prior), either by signalling the pure prediction error (Friston, 2005; Rao & Ballard, 1999) or the posterior probability (see (Aitchison & Lengyel, 2017)). As well as providing a mechanism by which Bayesian computations could occur, the theory offers a Bayesian reinterpretation of a number of phenomena such as repetition suppression, a consistent finding in which lower neural activation is seen on the second presentation of a stimulus. In the predictive coding framework, this activity attenuation is consistent with the idea that, having been viewed already, it is now 'predicted' (i.e. forms the prior), and as such gives rise to lower feedforward activity (Grill-Spector, Henson, & Martin, 2006; Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). Another neural response that follows naturally from the predictive coding theory is the 'oddball' response, whereby unexpected stimuli elicit a very high neural response: again this would be explained by the theory because unexpected stimuli cause a high prediction error (Lieder, Stephan, Daunizeau, Garrido, & Friston, 2013; Naatanen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). Nonetheless, although predictive coding is an intuitive mechanism for implementing Bayesian processes, the 'prior' in these studies is generally derived from horizontal connections, and feedback between layers for a concurrently perceived stimulus. This is quite far removed from the priors described in the empirical behaviour work above, which are reliant on explicit memory, or implicit knowledge of the statistical structure of the natural world. Thus it begs the question of where this more temporally-extended prior comes from, and how it is activated. We could argue that Hebbian learning – the long term strengthening of synaptic connections which may lead to easier communication between the neurons involved – is a natural answer to explain longer term priors, but it is unclear why one specific prior would be activated at a given moment. A full theory should account for the effects of different timescales of prior (c.f. (Hasson, Yang, Vallines, Heeger, & Rubin, 2008; Summerfield, Behrens, & Koechlin, 2011)). Finally, even if predictive coding could account for these questions (which it cannot in its current format), it is not the case that evidence for predictive coding is synonymous with evidence for a Bayesian mechanism, though it could be a useful starting point in outlining a mechanism by which Bayesian computations could be implemented (Aitchison & Lengyel, 2017).

Thus far in our quest to determine whether or not humans make 'good' decisions, we have seen a wealth of excellent empirical evidence that suggests that human behaviour in a host of decision making tasks generates the solution that would be derived from Bayes' theorem. Although our

examination of evidence that shows how Bayesian solutions may be realised in the brain suggested that the theory falls short the implementational level – there are no consistent accounts of how these complex and computationally demanding procedures are realised neurally – we nonetheless have consistent evidence for the 'optimality' of human choices, when 'optimality' is defined as making choices in line with the Bayesian solution.

Agents make irrational decisions in a host of situations

In this next section, however, I will outline a number of empirical findings that run counter to this notion. There are two central tenets implied by Bayes' theorem that will be a key focus here. The first is that all decision-relevant available evidence should be considered, with its reliability (variance) being the only factor that affects its impact on choice. The second is choice consistency: those who assert that humans behave in line with a Bayes optimal observer assume that it is only external noise that influences choices, but the internal decision process itself is consistently optimal. This means that were the same evidence to be presented again, the only thing that should lead to a different decision ought to be external noise in the process. However, a body of evidence – particularly from behavioural economics, but also as we shall see below from the field of perceptual decision making – has shown repeatedly that humans themselves are remarkably inconsistent, and display behaviours that challenge the notion that they are Bayes optimal.

The everyday work of a behaviour scientist is characterised by seeking evidence to support a given hypothesis. It is slightly ironic, then, that this 'confirmation bias' – the tendency to seek evidence to support, rather than disconfirm, a hypothesis – is considered irrational: only via falsification of all alternatives can a theory definitely be said to hold (Popper, 1959). Nonetheless it is a pattern of behaviour consistently shown by humans when making choices: rather than treating all evidence equally, evidence that confirms their pre-existing idea (whether this is derived from a bias, or from their initial impression when perusing the evidence) is given preference (Nickerson, 1998). Although this behaviour can in many situations be attributed to practicalities – as the example goes, to prove the theory that all swans are white, I must search the world for a black swan, which is generally impractically time consuming – impracticality has been shown not to be the only cause of this bias: even when definitive evidence can be acquired simply by falsification, people still by preference seek confirmatory evidence (Nickerson, 1998; Wason, 1968), and often overvalue it when it is found

(Bruner, Goodnow, & Austin, 1956; Doll, Hutchison, & Frank, 2011; Doll, Jacobs, Sanfey, & Frank, 2009). This asymmetric treatment of the available evidence presents a challenge to Bayesian theories of behaviour.

The decoy effect is linked to a more general category of preference reversal effect, whereby decisions change according to context. Again whilst often demonstrated in economic decisions (Tversky, Slovic, & Kahneman, 1990), this too has been shown in cognitive tasks: when choosing between two streams of rapidly presented numbers, participants' choices were affected by the information conveyed in a third, irrelevant distractor (Tsetsos, Chater, & Usher, 2012; Tsetsos, Usher, & Chater, 2010). A more specific example is the phantom decoy effect, a phenomenon in which, when several options are initially available, the removal of one of the options affects the relative ranking of those that remain. For example, after choosing a lemon gelato and hearing that the gelateria has run out, you opt for cremino al pistachio over the final option, coconut. The next week, upon hearing the lemon has been replaced with a tiramisu flavour, you order it only to find that this too has run out. This time when choosing between the remaining options (again cremino al pistachio and coconut), your second choice is coconut. This illustrates the phantom decoy effect: preferences between two options (coconut and pistachio) are dependent on a third, unavailable option (Pettibone & Wedell, 2000), and is part of a class of more general decoy effects whereby the preference of an irrelevant option affects decisions about independent options. Although decoy effects have again most prominently been studied in economic decisions (e.g. (J. Doyle, O'Connor, Reynolds, & Bottomley, 1999; Parrish, Evans, & Beran, 2015), they have also been demonstrated in perceptual decisions: when people had to choose the largest of three rectangles, the choice between the remaining two was influenced by the disappearance at choice time of the third option (Trueblood & Pettibone, 2015). As with confirmation bias, these contextual effects are not obviously in line with Bayesian processes (though see (Dayan & Solomon, 2010) for an explanation of how the response to irrelevant information could be interpreted in a Bayesian framework).

Although there are many others, the final non-Bayesian phenomenon to be considered here is anchoring. This highly consistent yet irrational behaviour involves making numerical judgements on the basis of recently seen, but entirely irrelevant, information: in the earliest experiments, Tversky and Kahneman spun a 'wheel of fortune' containing the numbers from 1-100. After indicating whether or not they believed that the number of African nations in the United Nations was greater or less than

the number spun by the wheel, participants then gave an estimate of the actual number of African nations. Despite the clear knowledge that the number the wheel spun was random, having watched it be generated, the second estimate was still significantly related to the wheel's number, as though it was acting as an "anchor" on estimates (Tversky & Kahneman, 1974a). The anchoring effect has been repeatedly demonstrated on a wide number of outcome measures (for example, how much an individual would pay for wine, the value they give to familiar products) with a wide range of arbitrary anchors (for example, digits in an individual's birthdate, social security numbers) (Ariely, 2009; Ariely, Loewenstein, & Prelec, 2003).

A new definition of 'optimal'

Thus we are left at a rather puzzling juncture. We began by reviewing clear evidence that humans make optimal choices, in line with a Bayesian observer. But we were then presented with evidence for the existence of a number of behaviours that humans consistently exhibit that fly in the face of Bayesian rationality. So how do we reconcile these conflicting patterns of results? Which is right?

The answer, I will argue, is neither: we are in fact using a limited definition of optimality. Bayes' theorem provides an excellent account of how humans make the best possible decisions given unavoidable external noise. However, I argue that a practical definition of an optimal decision process would be one that assumes that the best mechanisms operate in a way that can actively combat the effects of noise. Furthermore, Bayesian theories are excellently equipped to deal with noise arising from the senses during decision making, but overlook the influence of noise at other stages of the decision process. This short-sightedness explains why it is only under some circumstances that humans behave in line with the Bayes optimal solution. To come up with a more all-encompassing definition of 'optimality', we should acknowledge that noise affects decisions at multiple stages of the choice process, not just the earliest stage, and 'optimal' choices should overcome this corrupting source at all stages.

Furthermore, a theory of optimality should be biologically plausible. Not only (unlike the Bayesian account) should we have an idea about how it could occur at the implementational level, but this account should also explicitly acknowledge the fact that we are limited capacity beings. In the next sections I will put forward the idea that optimal decisions are those that *maximise the chances of a*

correct choice in the face of multiple sources of noise, via a plausible neural mechanism that explicitly takes into account our capacity limitation. I will show that this account, unlike the Bayesian theory, can provide a coherent explanation of the apparently contradictory findings just outlined.

Before I can start to do this, we need to understand in more detail the multiple sources of noise that can affect cognitive decisions, and particularly to understand the impact that *late noise* has on choices.

1.3. Multiple sources of noise

Imagine you are a teacher on a school trip to the Natural History museum, trying to herd the kids back onto the coach at the end of the day. You'd hoped that by dressing your class in fluorescent jackets they'd be easier to spot, but unfortunately all schools had the same idea, so you're working hard to pick out your class from the chattering mass in the entrance hall. In the terms I will be using here, this part of the task presents a challenge because of early, sensory noise (or uncertainty): children are all roughly the same size, they are all roaming round the entrance hall and they are all wearing the same clothes, so identifying the ones that go to your school is challenging. Once you spot them, you beckon them onto the coach, tracking how many of them you still need to find as you go. This second stage of the task presents a challenge because it is hard to keep count, particularly as the number of children steadily increases. I refer to uncertainty at this stage as "late noise", as it occurs after the initial sensory uncertainty. There is more than one source of late noise, this particular example illustrates 'integration noise'. It is important to note that all late noise is driven by internal rather than external sources: in this example, uncertainty is due to problems keeping count. Finally, you're aware that all but one child is safely on board, but in your exhaustion when the driver asks you whether it's time to go, you respond 'yes' on autopilot. Response variability caused by such factors is a form of late, response noise. As before, this source of late noise is internal (in this case, caused by tiredness). More formally, the dissociation between late and early noise can be defined as follows:

Early noise:

$$y = \sum_{i=1}^{n} f(X_i + N(0, \sigma))$$

Late noise:

$$y = \left[\sum_{i=1}^{n} f(X_i)\right] + N(0, \sigma)$$

in which y refers to the internal estimate of the available information, X_i refers to the sensory information X on sample i, N denotes noise drawn from a zero mean Gaussian distribution with variance σ, and n denotes the total number of samples viewed. A crucial distinction is that late noise, unlike early noise, is applied to the *summed* information.

The unique influence of late noise on decisions

What evidence is there then that this late noise has an important and unique impact (relative to early, sensory noise) on decisions? Research is limited, but investigations have shown that sensorimotor tasks are affected primarily by noise at a different stage of the information processing pipeline than that which influences cognitive tasks (Faisal & Wolpert, 2009; Hunt, 2014). More specifically, it has been pointed out that the types of cognitive tasks employed in the lab generally involve clearly perceptible stimuli (and as such are relatively unaffected by noise at the sensory processing stage), but are adversely affected by demand at the integration level, for example the need to integrate often highly conflicting pieces of evidence in a short time with a limited capacity system (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Rushworth, Kennerley, & Walton, 2005).

Part of the reason this research may be limited is because determining the independent influence of late noise is nontrivial: in fact, in many tasks it is impossible. For example in Brunton and colleagues' click count discrimination task (Brunton, Botvinick, & Brody, 2013), an identical pattern of behaviour could be caused by early noise in the absence of late noise (the noisy sensory perception of clicks, with lossless integration), or by late noise in the absence of early noise (noisy integration of perfectly detected clicks). In other words, in cases such as this, the impact of early and late noise is theoretically indistinguishable (Drugowitsch, Wyart, Devauchelle, & Koechlin, 2016).

Drugowitsch and colleagues thus conducted a series of experiments designed specifically to enable them to tease apart the effects of noise at different stages of the decision process. The task employed was simple: after viewing a series of oriented Gabor patches (see **figure 1.4**), participants were simply required to categorise the average orientation of the stream of stimuli. The length of the inference (number of stimuli to be integrated) and the dimensionality of the inference (number of orientation categories to choose from) varied independently between 2-16, and 2-3 respectively. If performance is mainly affected by early, sensory noise, choice variability should scale with the number of samples, because with each stimulus presented there will be sensory ambiguity about its precise orientation. Early noise also predicts that, when the log likelihood ratio of evidence in favour of each of the choices is calculated, this quantity will be correlated between the options, as sensory noise that favours one option disfavours the other(s). If performance is more affected by late, integration noise, choice variability should also scale with the number of samples because (as with counting children onto a school bus), cognitive demand increases as the number of stimuli to be integrated increases. However, this noise should have independent influences on the log likelihood of evidence favouring each sample, as it exerts its influence at the cumulative rather than momentary stage. Finally, if late selection noise is the key influence on decisions, choice variability should scale with the number of options, but not the number of stimuli (see **figure 1.3**). In all the tasks, the authors showed that choice variability was overwhelmingly better explained by a model that featured late, integration noise, and in a model comparison procedure in which all three sources of noise were included competitively, variance was again disproportionately explained by a late integration noise mechanism. This thus provides clear evidence of the unique influence of this source of variance on choice.

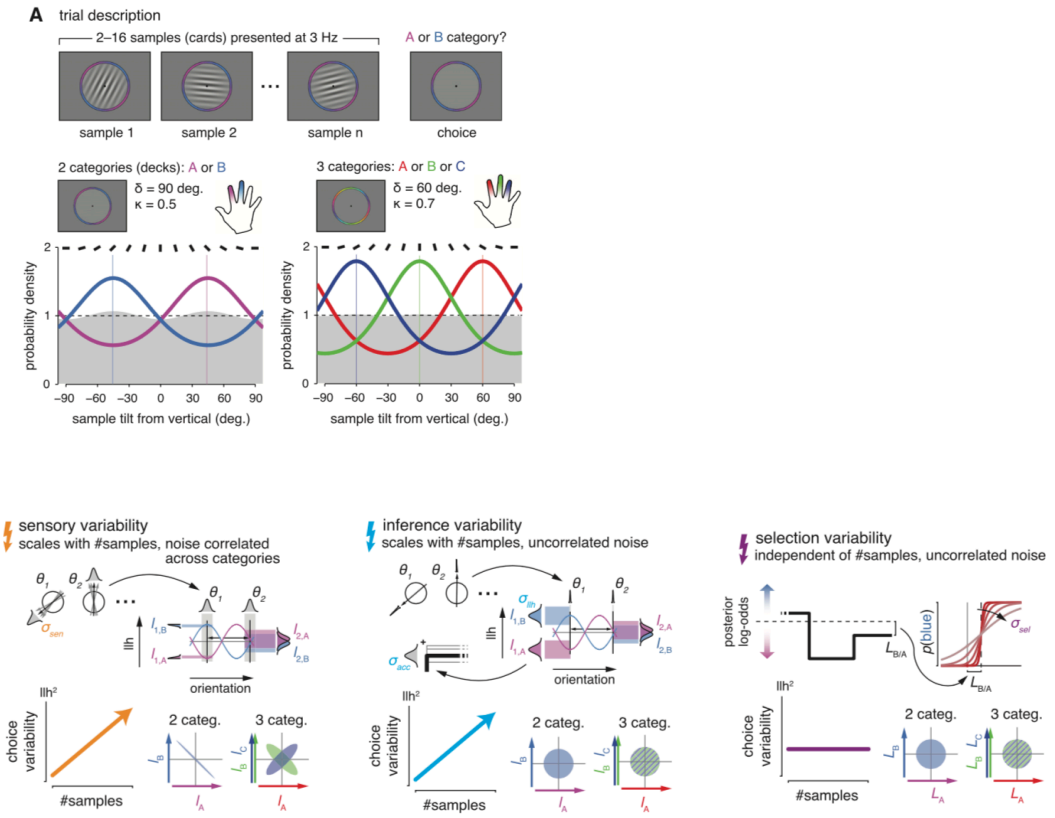

Figure 1.3: (Top panel) Illustration of the task used to tease apart the influences of early and late noise. The choice could involve two or three options, and was made on the basis of 2-16 samples. (Lower panels) Schematic of the effects different types of noise would be predicted to exert, as detailed on the figure. Figure taken from Drugowitsch, Wyart et al (2016).

Building on this notion, Herce Castañon and colleagues illustrated the dissociability of the influences of early and late noise on participants' confidence (Herce Castañon et al., 2018). Across a series of six experiments, observers were asked to judge on each trial the average orientation relative to a horizontal reference of eight simultaneously-presented Gabor patches. Sensory noise was manipulated by altering the contrast of the Gabors, making them more or less easy to perceive, whilst cognitive noise was manipulated by increasing the underlying variance of the angular distribution from which the eight samples were drawn, making the integration of disparate samples challenging. On 50% of trials participants were given a cue that was 75% valid, meaning that when they were more uncertain, participants should rely more heavily on the cue. The authors showed that participants did this when sensory, but not cognitive noise was higher. Furthermore, using two measures of confidence, the authors showed that participants appropriately calibrated their confidence as sensory, but not cognitive, noise increased. This study thus demonstrates the dissociability of the

influences of different sources of noise not just on performance, but on metacognitive processes such as confidence too.

Selective prioritisation of decision information is optimal under late noise

Having thus established that late noise has a unique influence on choices, we can now directly address the question of what strategy is optimal under circumstances of higher late noise, and whether or not people behave in line with this strategy. In what follows, we will show that the *selective prioritisation of decision information* – the down-weighting of certain pieces of information in favour of the prioritisation of others, via a number of different but related mechanisms – leads to the most accurate choices under conditions of late noise, and therefore provides a promising alternative definition for what it means to be 'optimal'.

Selective prioritisation via robust averaging

The selective prioritisation process is captured in a series of experiments that report the phenomenon of "robust averaging". In a low sensory noise task that involved reporting the average colour, shape, or orientation relative to a reference, of an array of eight clearly-visible stimuli, cognitive load was manipulated orthogonally by varying the strength of the evidence (mean colour or shape) as well as the variance of the distribution from which the stimuli were drawn. This latter manipulation in particular was linked to the level of late noise: the simultaneous integration of several highly disparate elements has been argued to lead to high integration noise, as noted above. The results from a regression analysis used to determine the impact of each stimulus on choice showed a striking effect of "robust averaging": information from "extreme", outlying stimuli carried less weight on the eventual choice (Li, Herce Castanon, Solomon, Vandormael, & Summerfield, 2017; Michael et al., 2015).

Robust averaging is counter to a Bayesian strategy: given that prior belief on each trial should have been equal for all alternatives (as each answer was equally likely to be correct), all stimuli ought to have been incorporated with respect to the information they conveyed. However the key question here is whether robust averaging nonetheless leads to good choices? The empirical answer is yes. In Li's study, the authors fit a power-law model via which stimulus values were transformed by an

exponent such that either inliers (close to the reference orientation) or outliers (extreme values, i.e. those that were empirically shown to be down-weighted) would have the greater impact on choice. The authors also estimated from each participant's data the slope of a logistic choice function that simulated late noise; higher values of slope indicated that the participant was relatively more affected by late noise. As expected from the qualitative results, human choices were best fit by an exponent value lower than 1, confirming that they over-weighted inliers in line with a robust averaging process, in particular when late noise was higher. However a second important insight from this model-fitting process was that robust averaging *improved* performance relative to no weighting, or the over-weighting of outliers, at all levels of late noise greater than zero. In other words, whilst robust averaging down-weights relevant choice information, this process of selective prioritisation of some information actually *improved* accuracy in the face of noise that occurs beyond the sensory stage.

What feature of this selective (robust averaging) mechanism may cause this to be the case? One potential answer is that, in these tasks, features (stimulus values, i.e. shape, colour, orientation) were drawn from a normal distribution in which the mode is always close to the category boundary (e.g. the red/blue boundary in the colour task, see **figure 1.1**) – i.e. the modal feature value is an 'inlier'. This means that through the process of robust averaging, decisions are optimally tuned to process information that is most likely to occur. Simulations for feature values drawn from uniform (rather than normal) distributions supported this hypothesis: under a uniform distribution, all feature values are equally likely to occur so robust averaging would lead to no optimal tuning, and the simulated data confirmed that asymmetric gain (as in robust averaging) conferred no advantage over linear gains on accuracy maximisation in this circumstance. Thus the crucial contribution of this work is to show not only that humans do engage in robust averaging, but also to demonstrate firstly that it is beneficial for accuracy under late noise, and secondly how it leads to this advantage.

Selective prioritisation of information via adaptive gain control

In these robust averaging tasks, all information was presented simultaneously and late noise was manipulated by increasing the variance of the underlying distribution from which the evidence was drawn. However, the majority of everyday decisions are temporally extended, meaning that evidence must be sampled repeatedly and these samples integrated across time, a process which also leads to late integration noise but of a slightly different nature. We see a similar pattern of behaviour performance in such tasks. Participants performing the same orientation classification task as that of

Li et al, but with the eight samples arriving sequentially rather than simultaneously, integrated evidence in line with its consistency with the running mean of information via an 'adaptive gain control' mechanism: discordant samples were integrated with lower gain than consistent evidence (Cheadle et al., 2014). As with robust averaging, this is a computationally efficient mechanism – not all information need be integrated fully, thereby saving precious neural processing resources – but nonetheless leads to accurate choices in the face of late noise.

Selective prioritisation of information in comparison tasks

In another temporally extended task, participants were asked to indicate which of two streams of numbers was higher on average. Each sample was presented sequentially, but simultaneously for the two streams: on each frame participants saw one sample of evidence from each stream. Participants again showed a pattern of selective weighting, this time prioritising the 'local winner' on each frame, i.e. the higher number of the two presented, and integrating it with enhanced gain. This process was captured by a model that features a bottleneck that discounts the gain of the weakest sample via a selective weighting parameter (see **figure 1.4**). As before, this selectivity was shown to lead to superior performance in the face of late noise (Tsetsos et al., 2012; Tsetsos, Chater, & Usher, 2015). Notably, this latter work implies a different process of selectivity: in the robust averaging and adaptive gain tasks, the inlying evidence closest to the running mean was preferentially integrated, whilst in the selective integration model, it is 'winning' evidence that gets preferential treatment. One potential determinant of which information is prioritised is the task context: in *simple categorisation* tasks, in which all information (presented serially or simultaneously) requires judgement in favour of one category or another, selective gain control prioritises that information that is most informative of this choice. In *comparison* tasks on the other hand, in which two (or more) streams of evidence each favour a different option, the more informative information is which stream is superior to the other – and as such, the locally winning information is selectively integrated to lead to the best choices. (See **chapter 5** for further discussion).

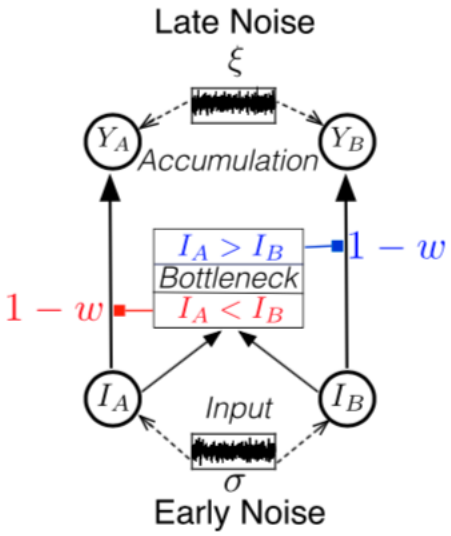

Figure 1.4: Illustration of the selective integration model. In an "A or B" categorisation task, I_A and I_B denote information in favour of category A and category B respectively. Information from the more prevalent category is integrated with full strength, whilst information from the less prevalent category is down-weighted before being corrupted by further late noise. Figure taken from Tsetsos et al (2015).

Selective prioritisation is neurally plausible and efficient

As alluded to above, all formats of the selective prioritisation process have the useful by-product of implying computationally efficient processing: some information need not be fully integrated, saving neural processing resources, and yet a good decision can still be reached. This is important given that we know humans (and other animals) have neural capacity limitations based on both cortical availability, which requires efficient coding in order to circumvent, as well as limited metabolic resources, which necessitates efficient computation (Franconeri, Alvarez, & Cavanagh, 2013; Lennie, 2003). Evidence from neural recordings suggests that selective prioritisation may fulfil both these criteria: during the simple categorisation task used by Cheadle and colleagues (involving the orientation classification of sequentially presented Gabor patches), differences were evident in the EEG signal: samples that were more similar to the running tally were encoded neurally with higher strength than those that were less similar, consistent with an adaptive gain control account that preferentially allocated resources to more consistent information. In blood oxygen level dependent (BOLD) signals in the parietal cortex, the dorsal medial prefrontal cortex, and the anterior insular cortex – all of which have been implicated in perceptual category judgements – there was a choice-

predictive encoding of the decision information provided by each sample that was significantly modulated by the difference in evidence between the current and previous samples (i.e. the discrepancy). This again supports the notion that there is a specialised processing mechanism that leads to enhanced processing of consistent evidence.

Divisive normalisation, a canonical neural process, provides a potential plausible mechanism via which information may naturally be integrated with greater strength depending on its position with respect to the rest of the information available: according to this framework, the responses of neurons are scaled by a common factor, generally based on the summed activity of a pool of neurons – in other words, information signalled by one neuron is relative to the information that other neurons are signalling too (Louie, Glimcher, & Webb, 2015; Rangel & Clithero, 2012). If this mechanism were able to be extended temporally such that differences between current and recent stimulation were also signalled, it could provide a good account of the underlying process that leads to selective integration in simple categorisation tasks.

In this section, we have seen that humans do not always behave in line with a Bayesian solution, but that they nonetheless often make accurate choices. Thus rather than defining "optimal" cognitive decisions as those that generate choices in line with the solution provided by Bayes' theorem, we could instead define "good" cognitive decisions as those that lead to correct choices in the face of late noise. If we use this definition, it seems that humans largely behave optimally. They do this via a process of selective prioritisation of some pieces of decision information over others, which itself is an efficient process. However, a decision is only good – no matter what your definition – if it is made *at the right time*. For example, if you spend several hours watching the weather forecast and make the decision to don wellie boots before leaving the house, it was the right decision 2 hours ago when it was still pouring, but now the sun is out and the boots are simply a sweaty inconvenience. All the decisions that we have described so far have been 'experimenter-paced' – participants were required to view as many stimuli as the experimenter determined, and make a choice at the allocated time – as is the overwhelming norm in the field. In the next section, we will discuss the ubiquitous situation in which a decision is made at a time of the agent's choosing. What factors predict *when* participants will commit to a decision, and is this timing appropriate?

1.4. Deciding when to decide

Agents must strike a delicate balance when determining *when* to make a choice. On the one hand, accumulating evidence over time is highly beneficial, as it allows them to generate a more precise estimate of the current evidence in order to inform choice: taking several samples of evidence allows an individual to accumulate them into an overall 'decision variable' – a single quantity that reflects the running combination of the evidence that has been seen so far – that has precision proportional to the square root of the number of samples taken. On the other hand, accumulating evidence is costly: delaying decisions can lead to lower reward in the long run, as fewer choices can be taken. This is particularly true in situations in which the correct answer remains elusive even after time has been spent generating evidence, as it is unclear whether continuing to sample will confer any extra benefit in such a situation – indeed, the explicit cost (in terms of possible reward) of continuing to sample increases as the decision duration lengthens (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012). Furthermore, the structure of the world can be volatile: as highlighted in our wellie boots example above, decisions must be executed at the appropriate time, otherwise the environment may change and that decision is no longer appropriate (Summerfield et al., 2011). Thus a crucial question in the decision sciences is: how do we decide when to decide?

In this section I will review the prominent models of decision timing, with a particular focus on the drift diffusion model, and outline behaviour and neural evidence that supports the idea that decisions are drawn to a close via the mechanisms proposed by the model. Finding that some elements of the process are not satisfactorily accounted for by this description, I will discuss extensions to the model that may improve its predictions, particularly under difficult conditions, where evidence for a decision is scarce. Finally, we will consider whether these models predict that humans draw choices to a close at an appropriate time.

The sequential probability ratio test and serial sampling models

One prominent answer to the question of *when* choices are made, first offered by Wald and Wolfowitz seventy years ago, is based on an extension of the likelihood ratio test that is foundational for signal detection theory (Green & Swets, 1966). When squinting in our attempts to determine whether the

animal in the distance is our loveable pet, or a fox to be chased away, signal detection theory decides on the basis of the likelihood ratio, by dividing the probability of the evidence given it being our pet, by the probability of the evidence given it being a fox (signal detection theory applies to momentary judgements in a static environment, and as such only considers one source of noise in deriving its solution). If we have no bias, a likelihood ratio higher than 1 should lead us to conclude that it was our pet, whilst a result lower than one suggests it is a fox. The principle of deriving the likelihood ratio in a static environment underpins the mechanism of the *sequential probability ratio test (SPRT)*: in circumstances that require the sampling of (conditionally independent) evidence over time, the optimal solution is derived by calculating the product of the likelihood ratio for each independent successively-occurring sample of evidence – or (equivalently) the sum of the log likelihood ratios – up to a given threshold. The SPRT allows us to determine the smallest number of samples required in order to make a decision with a given accuracy level (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Wald & Wolfowitz, 1949).

Whilst the SPRT is an exceptionally useful procedure, it also makes a rather implausible assumption: namely that observers are aware of the underlying distributions of evidence. Furthermore, it is not a particularly intuitive explanation of how humans (or animals) determine when to stop 'on the fly'. Hence the birth of serial sampling models (SSMs). The general format of SSMs is similar to that of the SPRT: evidence is accumulated over time in the form of a decision variable until this decision variable (DV) reaches a "bound", at which point a decision is triggered. In psychological terms, this bound may reflect our sense of certainty that our answer is correct, and we defer choices until that point. The frame in which the evidence is integrated is not specified in the original formulation of the model, thus the DV can be (as in the SPRT) defined as the log likelihood ratio. However many other formulations of the model (which are arguably more plausible) integrate evidence with reference to the format in which it is presented (for example, in an orientation discrimination task, the momentary evidence which is integrated would be the angular disparity of each stimulus from the reference) (Summerfield & Tsetsos, 2012). The bound reached (or in certain types of SSM, the identity of the evidence trace that hits the bound) determines the choice that is made. Thus such models are a parsimonious account of both *what* choice is made and *when* it is made.

Arguably the most prominent serial sampling model in the field is the drift diffusion model, DDM (see **figure 1.5**). This model assumes that the *relative* evidence in favour of two options is accumulated to

one of two bounds, and includes amongst its free parameters: the height of the decision-triggering bounds relative to the evidence starting point, the rate of drift (i.e. the average gain at each timepoint of the DV), and the level of noise (or variability) in the integration process. These parameters are estimated by fitting to the empirically observed reaction time distributions. In this model, the current 'state' of evidence in favour of each of the two options is modelled as a decision variable which drifts noisily towards one of two bounds (see figure). When the bound is breached, the accumulation process is terminated and the corresponding decision made (Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004).

The DDM is not the only sequential sampling model: other examples of such models differ most obviously in whether they assume that relative or absolute inputs are accumulated. Race models, for example, assume that evidence for each option is accumulated as an independent decision variable; these variables race towards their own boundary and whichever trace hits that boundary first determines the choice that is made. Race models have two clear advantages over models like the drift diffusion: firstly, it's intuitive to think how independent evidence traces could be represented neurally, and secondly, race models are able to represent more than two alternatives, and are therefore able to account for a much wider range of decisions (Bogacz, 2007; Furman & Wang, 2008; Tsetsos, Usher, & McClelland, 2011). However, in favour of the drift diffusion model, the influence of relative (rather than absolute) evidence on decisions is clear at all levels, from adjustments made due to background levels of light in sensory decisions (Baccus & Meister, 2002) to context dependent preferences in economic prospects (Tversky et al., 1990). This ubiquity perhaps explains the appeal of the relative model; either way, the drift diffusion model will be our key focus here due to its dominance in the field.

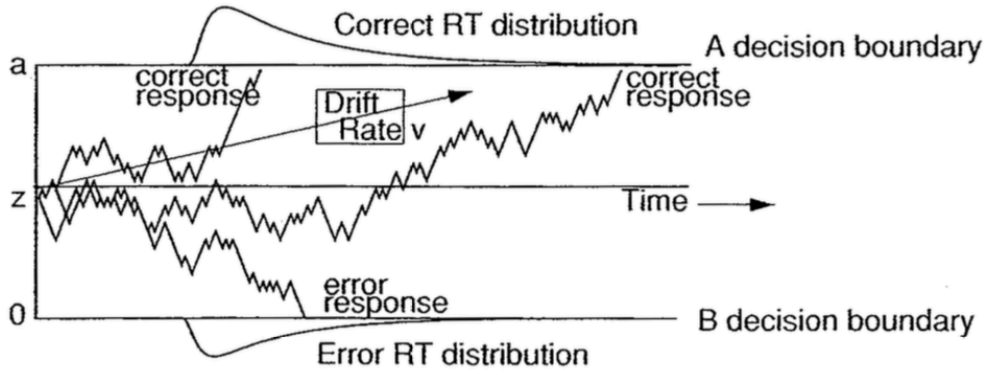

Figure 1.5: Illustration of the classic Drift Diffusion model (DDM) with three simulated paths. The starting point z is equidistant between the boundaries (0 and a) that reflect choice options A and B; shifting the starting point would lead to a bias towards one of the options. Figure taken from Ratcliff and McKoon (2008).

Behaviour support for the drift diffusion model as an account of human stopping times

What evidence is there then that the drift diffusion model is a good account of not just what people choose, but *when* they choose? Behaviour experiments have shown that it explains a host of features of reaction time data: the rightward skew of reaction time distributions arises naturally from the simple geometry of the model, as longer reaction times are associated with weaker traces (and therefore lower drift), meaning they cross the bound at more variable times. With a simple extension that involves variability in the drift rate or starting point of accumulation, the DDM can also captures empirically the longer reaction times seen in error trials (Ratcliff & Rouder, 1998). In the model, errors are made when the trace crosses the incorrect bound, and are also associated with lower drift. Thirdly, the model accounts for the speed accuracy trade off, as lower bounds cause speed to be emphasised, whilst higher bounds lead to greater accuracy but slower response times. In observed data, the general tendency for errors to be slower than correct trials is observed most frequently in the case that accuracy is being emphasised over speed; when the latter is emphasised, noisy traces are more likely to cross the incorrect bound quickly, thus the model accounts for this effect too. Finally, in order to account for decisions in which one decision is a priori more preferable, a bias can be introduced in the starting point of the decision variable in the model, such that the trace starts closer to the bound

representing the preferred option (Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004; Summerfield & Koechlin, 2008).

Neural support for the drift diffusion model from non-human primates

There is a prominent literature on how activity in the lateral intraparietal cortex (LIP) may reflect the predictions of the drift diffusion model. In the general form of the paradigm used, monkeys are trained to respond to a random dot motion (RDM) stimulus that contains motion information that tends left or rightward with varying degrees of coherence (the lower the coherence, the harder the task). Monkeys saccade towards a leftwards or rightwards target that reflects their choice about the net direction, and are rewarded with liquid for correct responses. During the experiment, the activity of a selection of LIP neurons that are spatially selective to the target locations – i.e. those that tend to fire when saccades are directed to the specific portion of space in which the targets are presented during the experiment – is recorded via extracellular single unit recording devices (see figure 1.6 and (Smith & Ratcliff, 2004) for an overview of this methodology). Using this method, some striking and consistent patterns of activity have been observed: on trials on which a correct response is made, the activity of neurons selective to the location of the correct target show a gradual ramping of activity during stimulus presentation, the gradient of which is greater when the stimulus coherence is higher. See **figure 1.6** for an illustration of how such a pattern is clearly reminiscent of the noisy drift towards a boundary of the DV in the DDM, which itself is predicted to be steeper on higher coherence trials. Furthermore, regardless of stimulus coherence and ramping strength, LIP firing rates are shown to reach a common level just before choice was executed, suggesting – as the DDM would predict – that decisions are terminated when a fixed 'threshold' is reached (Huk & Shadlen, 2005; Kiani, Hanks, & Shadlen, 2008; Kiani & Shadlen, 2009; Kim & Shadlen, 1999; Palmer, Huk, & Shadlen, 2005; Roitman & Shadlen, 2002). Furthermore, on choices in which there may be a strong prior preference towards one option over the other, there is evidence from single unit recordings that neural firing reflects this prior via increased baseline activity (Platt & Glimcher, 1999; Rorie, Gao, McClelland, & Newsome, 2010). When microstimulation is used to artificially inflate the drift, decisions are made faster, as would be expected if the decision variable is "boosted" closer to a bound. (Ditterich, Mazurek, & Shadlen, 2003). Thus this body of research provides a pleasingly intuitive demonstration of the potential neural implementation of the DDM, lending great support to the model as a good account of choices and stopping times.

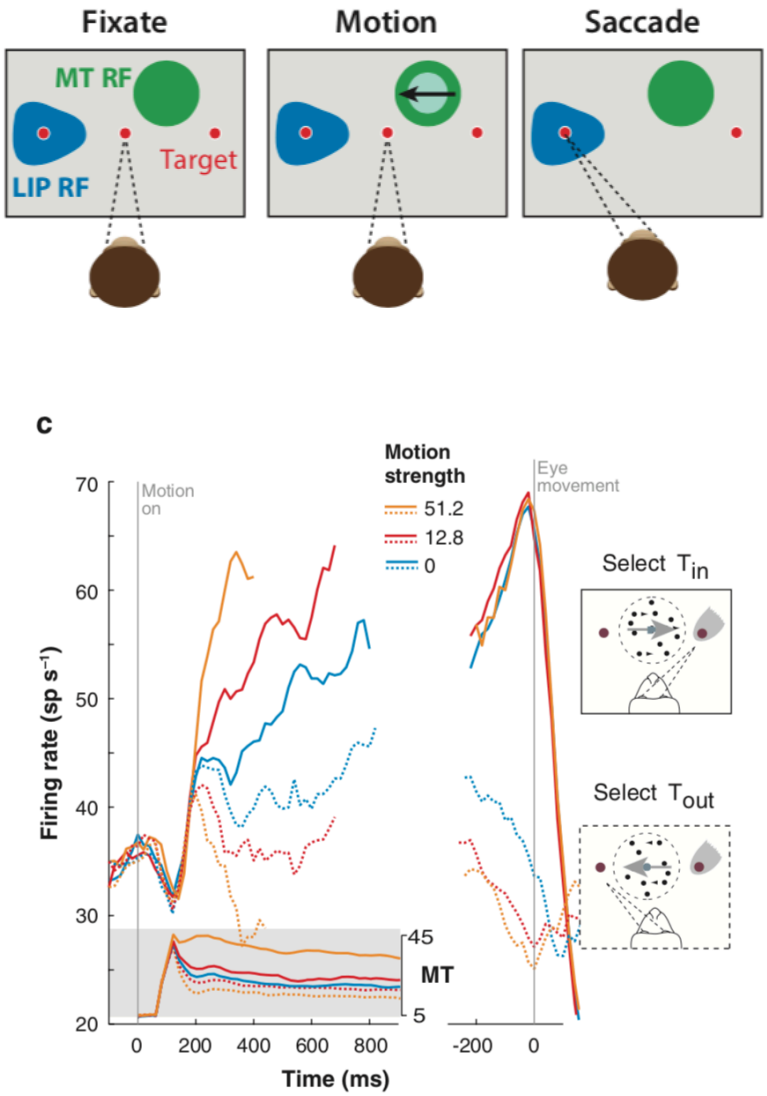

Figure 1.6: Illustration of the typical tasks used in the single unit recording paradigm (top panel, figure from Yates et al 2017) and the firing rates of LIP neurons (and medial temporal neurons in the grey panel) as a function of coherence level (lower panel, figure from Gold and Shadlen (2007)).

Challenges to the neural support for the drift diffusion model in non-human primates

However, perhaps all is not as it seems. It has been argued that the ramping signal in LIP could reflect attention and/or value, rather than the DV predicted by the DDM. Studies in experimental animals

often deliberately manipulate attention via reward contingencies associated with responses, and such studies rarely try to rule out the potential confound that the observed neural activity is simply due to attention to the to-be-rewarded choice (Maunsell, 2004). In the experiments described, the monkeys are highly proficient in the task, having completed thousands of training trials, and receive reward for correct answers. Thus it is certainly plausible that the neural response reflects an expected reward, rather than the disambiguation of a sensory stimulus which, due to extensive training, has long since ceased to be a challenge. Building on this idea, LIP neurons have been observed to correlate with the delay-discounted value of an offer shortly after its onset, before appearing to encode a choice in the build-up to action (Louie & Glimcher, 2010). This suggests that parietal regions are able to adapt rapidly and encode the changing subjective value associated with a choice (Sugrue, Corrado, & Newsome, 2004). It could be argued that encoding information in terms of its value rather than the sensory evidence does not negate the idea that the LIP reflects the decision variable: the DDM is a modality independent formulation of the evidence integration process and as such this drifting DV could reflect a range of different quantities, including the ramping relative probability of receiving a reward, arguably a highly relevant quantity for decision making.

A second element of the LIP activity that remains hard to explain how it is able to implement a 'bound': although firing reached a stereotyped level before the choice was executed, reminiscent of reaching a threshold, gradual ramping arguably isn't the type of activity we'd expect to see in a region associated with executing a binary choice, given that decisions are discrete rather than continuous. Furthermore, the fact that the activity rapidly attenuates after the decision is executed is indicative that the threshold is driven by input from other areas (Lo & Wang, 2006). This is not in itself necessarily incompatible with the notion that the pattern of activity reflects the decision variable, but for a full account it would be useful to understand which regions are involved in implementing the bound, and how their connectivity and communication may allow this to occur (more on this in relation to human data below).

Irrespective of whether we can accept this reasoning, there is a significantly more problematic challenge to the idea that this ramping neural activity reflects the DDM-predicted DV, which is that the LIP firing has been shown recently not to have causal relevance to the eventual choice or to its timing. When the LIP area was temporarily inactivated via the local application of GABA agonist muscimol, using the same paradigm, decisions about motion direction were still made with the same

accuracy, and evidence was accumulated at the same rate (Huk, Katz, & Yates, 2017; Katz, Yates, Pillow, & Huk, 2016). Importantly, this inactivation occurred in the very regions that displayed this characteristic ramping activity when not silenced, and nonetheless no impact was seen. Similarly in rats, silencing of the posterior parietal cortex actually enhanced task performance (Akrami, Kopec, Diamond, & Brody, 2018). Therefore this is a seriously problematic challenge for the theory that LIP implements an accumulation process as proposed by the DDM, and may weaken its ability to account for decision timing. Arguably, this challenge was long overdue: decades of work have shown that lesions to the parietal cortex lead not to disorders of decision making, but disorders of attention (Posner & Petersen, 1990); given that lesions are one of the few neural methodologies we use that allow us to make causal conclusions with relative confidence, we ought perhaps to have paid greater heed. It seems that researchers were perhaps distracted by the intuitive appeal of seeing a neural process that so closely resembles the model, and forgot that simply because some activity *looks like* a mechanism, it does not mean it is implementing that mechanism. To assume that it is involves attributing a semantic meaning to patterns of neural firing. To give an analogy, when I am watching a cat video on my computer, I would not expect the flow of electrical information across the motherboard to resemble the movements of the cat on the screen. Arguably Shadlen and colleagues' conclusions conflated the algorithmic and the implementational level of Marr's process description, leading to appealing but erroneous deductions.

Neural support for the drift diffusion model from human data

This does not mean that hope is lost for identifying neural evidence that supports the DDM: functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and magnetoencephalography (MEG) experiments in humans all open potential avenues. Focussing first on fMRI research, this paints an uncertain picture as there is not an established consensus of what would constitute evidence for the neural representation of a DV that evolves towards a bound. Arguably a higher BOLD signal is indicative of a harder choice, as more energy is required for the computation and the choice takes longer, meaning the sum of the firing will be higher. Such signals have been identified in the anterior insula and anterior frontal sulcus (Liu & Pleskac, 2011), and the medial prefrontal cortex (Pisauro, Fouragnan, Retzler, & Philiastides, 2017). However during a face/house categorisation task, Heekeren and colleagues found that the dorsolateral prefrontal cortex was *more* active for easier decisions, that it correlated with the signal difference between face and house selective regions (suggesting that the

comparison of outputs from different neural pools may drive the DV) and was linked to behavioural performance (Heekeren, Marrett, Bandettini, & Ungerleider, 2004). Given this opposite pattern of results, it is hard to draw firm conclusions about their meaning with confidence (see (Mulder, van Maanen, & Forstmann, 2014) for further discussion).

Furthermore, the DDM is a model of choice but is crucially simultaneously a model of reaction times, and neither of these experiments identify a link between the BOLD signal and decision termination. However, other findings may paint a clearer picture when it comes to trying to establish the neural basis of the decision bound. Information reaches the basal ganglia in the forebrain via excitatory or inhibitory pathways (Chevalier & Deniau, 1990; Redgrave, Prescott, & Gurney, 1999), and the basal ganglia receive information from LIP (Pare & Wurtz, 2001), meaning they may be well placed for a role in decision execution via the implementation of a bound (Lo & Wang, 2006), a notion exacerbated by the fact that in the striatum, D1 and D2 dopamine receptor uptake is associated with positive and negative sensorimotor reinforcement respectively (Doll & Frank, 2009; Frank, 2005; Frank, Seeberger, & O'Reilly R, 2004). Thus although evidence for the correlate of the DV itself is perhaps lacking, we arguably have more convincing evidence for the potential neural location of decision termination.

Event related desynchronisation is the transient dampening or blocking of rhythms within a given frequency band that can be observed in electro- and magneto-encephalographic (EEG and MEG) data. Beta band frequency desynchronisation in parietal regions has been linked particularly to cognitive (rather than sensory) processing (Pfurtscheller, Neuper, & Mohl, 1994). Its link to cognitive processes has been suggested to take the form of involvement in the evidence integration process (Donner, Siegel, Fries, & Engel, 2009), based partly on work that has shown that beta power decrease in the hemisphere contralateral to the response effector is linked to motor response preparation (L. M. Doyle, Yarrow, & Brown, 2005). In EEG, scalp activity has been shown to ramp proportionally to the extent of sensory evidence, akin to a DDM decision variable (Philiastides, Heekeren, & Sajda, 2014). Focusing on specific components, the centro-parietal positivity component (CPP) shows a ramping profile, scaled by difficulty level, over the course of a perceptual decision making trial, and the peak of the component predicts response times (Kelly & O'Connell, 2013; O'Connell, Dockree, & Kelly, 2012). Predicting response time is important given the focus of our current discussion, yet the fact that it can't simultaneously predict choices themselves is problematic to a quantity that is being viewed as akin to a DDM decision variable, which does both. Some have argued that this suggests that

the CPP instead reflects a choice-independent quantity such as confidence (Urai & Pfeffer, 2014). Ultimately, then, the neural evidence for the quantities predicted by the DDM exists but remains ambiguous in both humans and non-human primates.

Does the drift diffusion model make choices at the appropriate time?

A remaining pertinent question is: does the DDM make decisions at a 'good' time? Even if neural evidence is somewhat lacking, we know it accounts well for behaviour, so are these choices appropriately timed? Like the SPRT, the DDM is able to produce the fastest decisions for a given level of accuracy, and is furthermore able to maximise desirable values better than other similar sequential sampling models (Bogacz et al., 2006). For example, Bayes risk (BR) is a cost function that is the weighted sum of decision time and error rate:

$$Bayes' \; Risk = c1 \cdot Decision \; Time + c2 \cdot Error \; Rate$$

where c1 is a time cost, and c2 is a cost for errors (which, given that the two types of cost are comparable, must be scaled by time in some way). These two cost parameters can either be fit to the data, or established via a prior based on the relative importance in a given task of speed versus accuracy. Another way of assessing value over time is via the reward rate (RR): the proportion of correct trials divided by the average duration between decisions (see e.g. (Gold & Shadlen, 2002)). These two functions – Bayes' risk and the reward rate – differ in their penalties – BR has an explicit cost for time, RR just assumes that long times are penalising in that they slow the reward rate. They also have a different emphasis on speed/accuracy - RR is unconcerned with getting the answer right, as long if it means that overall the reward is still high. BR on the other had has an explicit cost to incorrect choices, independent of the time cost. However, whichever definition is used, there will always be a set of parameters for the DDM that maximises the value given more effectively than other similar sequential sampling models (Bogacz et al., 2006). As such, we know that the DDM certainly has the potential to lead to appropriately-timed decisions. However, importantly, simply because there exists an optimal set of parameters does not imply that humans (or monkeys) are able to identify them.

Urgency signals in the drift diffusion model

Furthermore, a specific area of contention is whether the DDM leads to appropriately-timed choices in particular when the decisions are hard. Consider the fable of Buridan's ass: the hungry, thirsty animal is equidistant between sources of food and water. Given the equally strong appeal of both options, the ass is unable to choose where to go first, and therefore perishes. In everyday life there are decisions that even after a significant amount of evidence has been gathered remain difficult, either because no option is clearly correct, or because all options seem equally likely. In the original form of the DDM, ambiguous evidence would drift between the two boundaries for a prolonged period, unduly delaying the choice to be made. So how do agents bring decisions to a close when, even after collecting evidence, the answer is unclear? There is a reasonable amount of evidence that suggests that humans implicitly track the time elapsed during evidence accumulation. Intuitively, if it has taken us double the time than it took on a different occasion to reach the same level of evidence in favour of a decision, it suggests that the source of evidence this time round is noisier – the ratio of signal to noise is lower – and therefore our confidence should be lower. Having used this information to deduce that we are in a low signal to noise environment, we can further reason that further accumulation will likely confer little benefit in identifying the best choice – so it is better to curtail such decisions and move on. People seem to behave in line with this reasoning (Hanks, Mazurek, Kiani, Hopp, & Shadlen, 2011; Kiani, Corthell, & Shadlen, 2014; Zylberberg, Fetsch, & Shadlen, 2016).

A more explicit mechanism that has been proposed via which agents may track the passage of time and draw decisions to a close is an 'urgency' signal: an evidence-independent signal that increases as the accumulation process continues, naturally drawing ambiguous decisions to a close. In the framing of the drift diffusion model, this could occur via a factor added to the drift, leading to steeper gain and an eventual boundary crossing even when evidence is weak, or via a collapsing bound, such that over the course of time, increasingly less evidence is required to trigger a choice. Evidence for such a signal has been widely supported via a number of avenues: model-based behaviour evidence from Thura, Cisek and colleagues supports their variant of a serial sampling model that explicitly incorporates an urgency signal (Thura, Beauregard-Racine, Fradet, & Cisek, 2012; Thura & Cisek, 2014), a notion that is in line with Drugowitsch and colleagues who concluded that reaction times under various levels of task difficulty could only be explained by a model with an urgency signal (Drugowitsch, DeAngelis, Angelaki, & Pouget, 2015). Neurally, ramping activity in single neurons in

the absence of any coherent signal in the stimulus has been suggested to reflect an urgency signal (Churchland, Kiani, & Shadlen, 2008). An odour discrimination task in humans during fMRI combined behaviour and neural evidence, showing that the activity in the OFC and behaviour data together were best accounted for by a drift diffusion model with collapsing bounds (Bowman, Kording, & Gottfried, 2012). These are but a few examples of a wealth of research that has supported the incorporation of an urgency signal into models that account for choice reaction times.

However, two recent reviews of such experiments concluded that overall, reaction times were best fit by a fixed bound model (Hawkins, Forstmann, Wagenmakers, Ratcliff, & Brown, 2015; Voskuilen, Ratcliff, & Smith, 2016). However, these reviews arguably did not appropriately consider a crucial factor: urgency signals should only be obvious in tasks in which the decision is difficult, or where the decision difficulty varies considerably between trials and it is therefore unclear before the trial onset at what level to set the bounds. If the tasks are easy, the DDM predicts that the bound will be reached anyway, before such time as an urgency signal is apparent. In other words, the conclusion that an urgency signal doesn't exist may have been drawn simply because in the experiments that were reviewed by Hawkins and colleagues, the tasks were too easy. This notion is supported directly by a study that compared easy and hard versions of a task directly, which showed that indeed, collapsing bound models are evident only when the task difficulty varies from trial to trial, as would be expected (Malhotra, Leslie, Ludwig, & Bogacz, 2017). Ultimately, it seems reasonable to conclude that, in the absence of a clear choice, some factor must drive decisions to a close in order to prevent endless deliberation. Taken together, the evidence suggests that a time dependent 'urgency' signal, evidence for which has been seen behaviourally and neurally, is the most plausible explanation.

Over the course of this section, a thorough examination of the evidence supporting the most prominent model of human stopping times – the drift diffusion model – has shown that it provides an excellent fit to human reaction time distributions, with the incorporation of an 'urgency' signal to bring decisions to a close when the evidence is ambiguous. Assuming agents are able to determine the appropriate parameters, it leads to the best-timed choices under a number of formulations. Neural evidence in favour of the model, whilst existent, is less easy to interpret.

1.5. Chapter summary and thesis outline

I began this chapter by outlining the wealth of evidence that suggests that, when making simple decisions, humans reliability-weight evidence just as an optimal observer should. However, I also noted that traditional accounts of human optimality lack a plausible mechanism by which they could be implemented. In **chapter 2,** I present empirical evidence that humans reliability-weight information in a simple, multi-sample categorisation task. Using data from human behaviour as well as concurrent EEG recording, I provide evidence that this pattern of results is accounted for by a simple and biologically plausible mechanism.

I then went on to note that the traditional definition of 'optimal decisions' – those that reliability-weight the information in the manner of a Bayesian observer – may be an appropriate definition only in the simplest of tasks. When choices are corrupted by late noise – noise that arises from internal processes rather than external sensory sources, such as during the integration of multiple pieces of evidence – "optimal" choices may be better defined as those that are most likely lead to correct choices in the face of late noise. I detailed preliminary evidence that such choices involve the selective integration of decision information. In **chapter 3**, I provide empirical evidence that shows that humans selectively integrate decision information during a cognitively challenging multi-sample task. I show that this behaviour can be captured by a selective integration model, the predictions of which are manifest in the concurrent EEG data.

Finally, I discussed factors that drive decisions to a close. In **chapter 4** I describe a series of three experiments in which participants made decisions in a self-paced manner. I show that a model that integrates evidence selectively until a threshold that collapses over time is reached provides the best explanation of human reaction times. I show that the predictions of this model are again evident in the concurrently-recorded EEG data.

The thesis is concluded in **chapter 5** with a more general discussion of the common themes of the work from the three empirical chapters, during which I highlight questions that remain unanswered, and present ideas for future research directions.

2. Near-optimal integration of magnitude in the human parietal cortex

Humans are often observed to make optimal sensorimotor decisions, but to be poor judges of situations involving explicit estimation of magnitudes or numerical quantities. For example, when drawing conclusions from data, humans tend to neglect the size of the sample from which it was collected. Here, we asked whether this sample size neglect is a general property of human decisions, and investigated its neural implementation. Participants viewed eight discrete visual arrays (samples) depicting variable numbers of blue and pink balls. They then judged whether the samples were being drawn from an urn in which blue or pink predominated. A participant who neglects the sample size will integrate the ratio of balls on each array, giving equal weight to each sample. However, we found that human behaviour resembled that of a near-optimal observer, giving more credence to larger sample sizes. Recording scalp electroencephalographic (EEG) signals whilst participants performed the task allowed us to assess the decision information that was computed during integration. We found that neural signals over the posterior cortex following each sample correlated first with the sample size, and then with the difference in the number of balls in either category. Moreover, lateralized beta-band activity over motor cortex was predicted by the cumulative difference in number of balls in each category. Together, these findings suggest that humans achieve statistically near-optimal decisions by adding up the difference in evidence on each sample, and imply that sample size neglect is not a general feature of human decision-making.

2.1 Introduction

Decisions often involve integration of evidence from multiple sources. We discussed extensively in chapter one the idea that optimal choices in simple decisions can be made when information is weighted by the trustworthiness (or *reliability*) of each source. When human sensorimotor behaviour is refined through experience, it often resembles that of an ideal observer (Kording, 2007; Pouget et al., 2013). For example, humans pointing towards a target assign more weight to prior knowledge about its location when sensory evidence is indistinct, as an ideal observer should (Kording & Wolpert, 2004). Humans and monkeys can learn to weight a train of symbolic cues in direct proportion to the informativeness with which they predict a rewarded response (Gould, Nobre, Wyart, & Rushworth, 2012; Wyart, de Gardelle, Scholl, & Summerfield, 2012; Yang & Shadlen, 2007). When visual and haptic cues offer potentially conflicting information about the size of an object, visual information is less influential when corrupted by noise (Ernst & Banks, 2002). When using information from product reviews to update their initial impression of the product, participants update more strongly when the number of people who contributed to the reviews was larger, and was therefore a more reliable source of information (De Martino et al., 2017). These findings have prompted the claim that humans and monkeys have evolved to make optimal decisions, i.e. those that account for the relative uncertainty associated with each source of choice-relevant information.

However, human choices are not always optimal. For example, when asked to evaluate hypothetical scenarios involving numerical magnitudes, humans often make biased or inaccurate responses (Griffin & Tversky, 1992). Consider the problem of estimating whether men outnumber women on a university degree course. The approximate male:female ratio from a small seminar group is a less reliable estimator than that from a large lecture class, because our confidence in an estimate should be determined by its standard error, which is inversely related to n, i.e. to the sample size. Given both observations (small seminar versus large lecture), an optimal solution to this problem – which is given by combining the binomial probability associated with each sample – will afford the larger sample more weight; simply averaging the two ratios may lead to a biased decision. However, when confronted with problems of this nature, humans can be excessively reliant on the relative balance of evidence, overlooking the information about sample size and drawing erroneous conclusions from data (Griffin & Tversky, 1992; Tversky & Kahneman, 1974b). Instances of this suboptimality are not

limited to binomial estimation problems: when asked to judge whether the average height of a group of humans exceeds a fixed value, humans disregard whether the group is composed of 10, 100 or 1000 individuals. Even researchers who use statistics regularly to evaluate data have been observed to display this 'sample size neglect' (Tversky & Kahneman, 1971).

Why, then, do humans account for the reliability of information in some situations and not in others? One possibility is that sample size neglect might not be a ubiquitous feature of decisions, but instead depends on the format in which the decision information is provided. For example, sample size neglect might occur when decision problems are presented as descriptions of hypothetical scenarios, but not when participants learn to make decisions in an experience-dependent fashion, i.e. through feedback that reveals whether a classification judgment was correct or incorrect (Hertwig & Erev, 2009).

Here, we asked whether humans performing a psychophysical task display sample size neglect, or whether they integrate information about numerical magnitudes optimally. Our task was an expanded judgment task, variants of which have previously been used to interrogate information integration during perceptual decision-making (Wyart, de Gardelle, et al., 2012; Yang & Shadlen, 2007). Our approach thus investigates sample size neglect via an experimental framework that has been widely used to understand the neural and computational mechanisms by which perceptual inputs are integrated and categorised (Gold & Shadlen, 2007). Observers viewed a succession of eight discrete visual events ('samples') in which a variable number of pink and blue balls were displayed, and subsequently decided whether they had been drawn from a larger pool of predominantly pink or predominantly blue balls. Our initial question was whether humans gave more credence to samples that contained more balls. For example, consider two different samples each offering a 2:1 ratio of blue:pink balls, one with a total of 3 balls (2 blue, 1 pink) and one with a total of 12 balls (8 blue, 4 pink). If humans exhibit a sample size bias, their choices would reflect the integration of these ratios of evidence (i.e. they would weight samples with 3 and 12 balls equally). However, if humans are optimal, they will give more weight to samples with more balls.

We initially predicted, in line with the second body of research outlined above, that humans would show sample-size neglect during integration of magnitude information. However, we found strong support for the opposing hypothesis: human choices were 'near-optimal', in that their choices resembled those that might be made by an ideal observer (in this case, one who was calculating the

binomial probability of the dots on each sample being drawn from one urn as opposed to the other). In the light of this finding, we turned our attention to understanding how this behaviour might be achieved at the neural and computational level. In our task, two simple strategies would allow participants to arrive at the near-optimal solution. Firstly, participants could add up the *absolute* evidence for either response, by integrating the blue and pink dots independently and comparing the resulting tallies (we call this the tally model; it is related to the 'race' model of perceptual decision-making, (Smith & Vickers, 1988; Vickers, 1979)). Alternatively, participants could add up the *relative* evidence for either choice, by integrating the difference in the number of blue or pink dots on each sample (we call this the difference model; it is related to the drift-diffusion model of perceptual choice, (Ratcliff & McKoon, 2008)). Because the behavioural data did not allow us to arbitrate among these possibilities, we turned to neural recordings, and measured the scalp electroencephalogram (EEG) whilst participants performed the task. Although we observed a correlate of the absolute number of pink or blue dots over posterior electrode sites, this was rapidly followed by a correlate of their relative difference. Over centro-parietal sites previously implicated in perceptual categorisation, we observed stronger correlates of both the momentary and integrated difference signal. Together, these findings suggest that humans solved the task by accumulating the relative difference in magnitudes of evidence on each sample, allowing them to perform the task near-optimally via a computationally tractable strategy.

2.2 Methods

Participants

Fifty-four participants (31 female, 23 male) were recruited. All reported normal or corrected-to-normal vision and no history of neurological or psychiatric disorders, and gave written informed consent in accordance with local ethical guidelines. Participants for the behavioural pilot (n = 15) and control experiment (n = 19) received £10 in compensation, whereas those undergoing EEG (n = 20) received £25. Data from 4 EEG participants were excluded (prior to preprocessing) due to excessive movement and/or electrical interference, leaving n = 16 for that experiment. For behavioural analyses, we included all pilot and valid EEG participants (total n=31).

Task design and stimuli

In both the behavioural and EEG experiments, participants completed a probabilistic decision-making task (**Figure 2.1**). On each trial, one of two virtual "urns" was pseudorandomly selected: either one with a 60:40 predominance of pink balls (50% trials), or one with a 60:40 predominance of blue balls (50% trials). Eight draws (with replacement) of 2,4,6,8,10 or 12 balls were made from the relevant urn. The blue and pink balls drawn were represented as dots in a circular aperture on each of the eight sample screens for each trial. After viewing the 8 samples, participants indicated whether the samples were drawn from the predominantly pink urn or predominantly blue urn.

The visual stimuli were presented using the Psychophysics-3 Toolbox (Brainard, 1997) running in MATLAB (The Mathworks, Natick, MA) on a 17' CRT monitor with resolution of 1024x768 pixels and a refresh rate of 60Hz. All stimuli were presented on a black background. Participants viewed the stimuli in a quiet darkened room approximately 70cm from the screen. Before the experiment began, instructions were presented to participants on screen, including a visual cue indicating the ratio of pink:blue balls in the two urns. Each trial began with a white central fixation cross with lines of length 60 pixels (for 1000 ms), followed by a blank screen (1000 ms). Subsequently, participants viewed a forward mask, eight sequentially occurring sample screens, and a backward mask, each occurring for

150 ms with 150 ms inter-stimulus interval (ISI). Each sample screen consisted of an array of blue, pink and white dots, each 20 pixels in diameter, randomly spatially distributed (minimum separation 10 pixels) among 71 possible locations within a circular aperture of 300 pixels diameter. The number of blue and pink dots was determined by randomly drawing between 2 and 12 balls from one of two virtual urns as described above. White balls were distracters, which served to decorrelate decision information from low-level visual input; 1-6 white balls were randomly added to each sample. The mask screens were identical to sample screens except that the coloured balls were green and orange, and participants were instructed to ignore them. These screens helped ensure that participants were not unduly swayed by the first or last sample (i.e. avoided perceptual primacy/recency effects).

Following the presentation of the samples, participants indicated via a key press whether they thought the balls on that trial were drawn from the mainly pink urn (m key, with right hand), or from the mainly blue urn (z key, with left hand). Auditory feedback was given with a high tone (800 Hz) for correct answers and a low tone (400 Hz) for incorrect answers. Before they began the task, participants were given clear instructions, a visual representation of the decision problem and urns, and 2 practice trials with feedback.

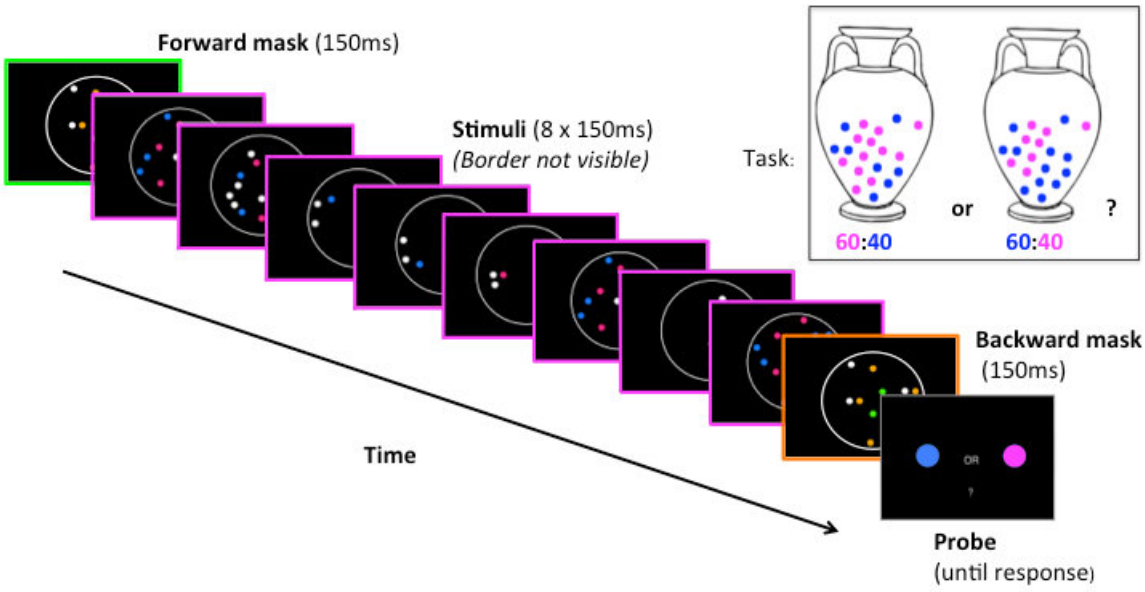

Figure 2.1. Experimental protocol. Each trial commenced with a blank screen for 1 second, followed by a central fixation cross for 1 second. Eight draws (with replacement) of 2, 4, 6, 8, 10 or 12 balls were made from a virtual "urn", either with a 60:40 predominance of blue balls (50% of trial) or with a 60:40 predominance of pink balls (50% of trials), and each draw was represented on screen as coloured dots within a circular aperture. Each screen also contained between one and six white distractor dots and was displayed for 150ms with 150ms interstimulus interval (ISI). These target screens were preceded and followed by a forward and backward mask created in an identical fashion, except that the coloured dots were orange and green, and participants were instructed to ignore them. At the end of the sample series, participants saw a screen prompting them to respond with a keyboard press. Feedback was given on each trial, with a high-pitched (800Hz) tone for correct response and a low-pitched (400Hz) tone for an incorrect response.

The behavioural pilot differed from the EEG experiment in three ways: (i) the presentation time and ISI were 250ms each rather than 150ms, (ii) each participant completed 288 rather than 570 trials, and (iii) the forward and backward masks were omitted.

Statistically optimal solution

Let us denote the number blue and pink balls on sample k as $d1_k$ and $d2_k$ respectively, with $n_k = d1_k+d2_k$. The probability of drawing a blue ball from urn one (predominantly blue) was p and the

probability of drawing a blue ball from urn two (predominantly pink) was 1-*p*. The converse was true for pink balls. By design, *p* = 0.6.

The statistically optimal solution to the task is given by the integration of binomial probabilities. For any sample *k*, the likelihood that the balls are drawn from the blue urn is given by

$$p'_k = B(d1_k, n_k, p)$$

where B denotes the binomial distribution. The optimal decision rule is defined by whether the sum of the log likelihood ratios of each sample coming from either urn is greater or less than zero. On any sample *k,* the evidence or optimal *decision update* DU for or against each response can be quantified as the log-likelihood ratio:

$$log\left(\frac{B(d1_k,n_k,p)}{B(d1_k,n_k,1-p)}\right)$$

Models of evidence integration

We considered three models of evidence integration and choice that human observers could be using to solve the task: a suboptimal model, and two models that arrived at the statistically optimal solution via two qualitatively different computations. Thus, these latter two make identical behavioural predictions, but different predictions about the neural activity that would accompany each sample. We first defined the suboptimal model, one in which evidence was not weighted by sample size. In the *ratio model*, the momentary decision update, $DU_{k,\text{ratio}}$, was based on the log ratio of blue to pink dots:

$$DU_{k,ratio} = log\left(\frac{d1_k}{d2_k}\right)$$

Thus, the ratio model ignores the sample size: for example, the same value of DU is obtained for $d1_k$ = 1 and $d2_k$ = 3 (where the number of balls n_k = 4) as for $d1_k$ = 3 and $d2_k$ = 9 (where n_k = 12). Choices were made according to whether the decision variable DV_{ratio}, was greater or less than zero:

$$DV_{ratio} = \sum_{k=1}^{8} DU_{k,ratio}$$

This policy accounts for the behaviour of participants in the experiments of Tversky and colleagues described above.

Next, we considered two models that are formally equivalent to the statistically optimal solution, but that solve the task via more biologically plausible mechanisms. We call these the *difference* and *tally* models. The *difference model* computes its decision update (DU) by taking the difference between the number of blue and pink dots on each sample, and adds up these differences to form the decision variable (DV). Choices are then made according to whether this DV is greater or less than zero:

$$DU_{difference} = d1_k - d2_k$$

$$DV_{difference} = \sum_{k=1}^{8} DU_{k,difference}$$

This model predicts that brain signals accompanying each sample should correlate with the difference between the number of pink dots and the number of blue dots, i.e. $d1_k$-$d2_k$.

By contrast, the *tally model* adds up the number of blue and pink dots in each stream without computing their difference. This model thus computes two momentary decision variables for each sample, and the choice is made according to whether the sum of blue dots (DU*1*) exceeds the sum of pink dots (DU*2*) or vice versa (i.e. according to the sign of the DV):

$$DU1_{tally} = d1_k$$
$$DU2_{tally} = d2_k$$

$$DV_{tally} = \sum_{k=1}^{8} DU1_{tally} - \sum_{k=1}^{8} DU2_{tally}$$

This model predicts that brain signals accompanying each sample should correlate with the total number of blue and the total number of pink dots, i.e. $d1_k$ and $d2_k$.

Thus, whilst the quantity DV on which decisions are based is identical for the difference and tally models, because they arrive at this decision variable via different computations, they make different predictions about the neural activity that will accompany each sample. The difference and tally models are related (but not identical) to the *diffusion* and *race* models often used to model reaction times in psychophysical tasks (Ratcliff & McKoon, 2008; Smith & Vickers, 1989; Vickers, 1979).

For comparison with human data, DV_{model} was corrupted with values drawn from a Gaussian distribution, with mean of zero and a standard deviation of σ, before being used to generate categorical (model) choices. The noise parameter σ was fit to group performance separately for the tally and difference model versus the ratio model, but yielded very similar values (in log units: 0.51 for the ratio model, 0.59 for the tally and difference models). Note that varying this parameter simply changed overall model performance without affecting the qualitative pattern of results. The distributions of DV associated with the tally and difference models were very similar; each was roughly normally distributed and ranged from -3 to 3 in log units.

Behavioural analyses

We compared human and model performance in a number of ways. First, we plotted psychometric functions to envisage how the probability of responding pink, p(pink), varied as a function of the decision variables (DVs) predicted by the ratio, difference and tally models (each binned into deciles). Second, we used a probit regression model to estimate the influence that each of the sample positions (from first to last in the sequence) had on choice, and plotted how p(pink) was predicted by the number of blue and pink balls on each of the 8 samples, as follows:

$$p(pink) = \phi\left[b + \sum_{k=1}^{8}(w_k \cdot d1_k + w_k \cdot d2_k)\right]$$

Where ϕ denotes the standard Normal cumulative distribution function.

Finally, we tested how the influence of each sample on choice differed as a function of the total number of dots, $d1_k + d2_k = n_k$. The ratio model predicts that all samples should have the same weight, irrespective of n_k. The difference and tally models predict that the weight carried by each sample should increase as a function of n_k, in line with the statistically optimal (binomial) process. On each trial, we sorted the 8 samples according to n_k, from lowest to highest, denoting the rank of each sample *j*, and used probit regression to calculate the coefficients that best mapped DU_{ratio} onto the choices made by humans, and the various computational models:

$$p(pink) = \phi\left[b + \sum_{j=1}^{8} w_j \cdot DU_{ratio,j}\right]$$

The logic of this analysis is that if humans are integrating the ratio of evidence, then the resulting coefficients should be flat across different values of n_k, (i.e. over the ranks *j*) whereas if they are performing (near-) optimally, then the coefficients should grow with *j.*

EEG acquisition and preprocessing

A Neuroscan EEG system with NuAmps digital amplifiers was used to record EEG signals from 32 Ag/AgCl electrodes, located at FP1, FPz, FP2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, POz, O1, Oz and O2, plus 4 additional electrodes used in a bipolar montage as horizontal and vertical electro-oculograms (EOGs) and two electrodes located at the mastoids used as reference. The electrode impedances were kept below 10KΩ. EEG signals were recorded at a sampling rate of 1kHz and high-pass filtered online at 0.1Hz.

Preprocessing was carried out using the EEGLAB toolbox for MATLAB (Delorme & Makeig, 2004) and custom scripts. The data were downsampled to 250Hz, and epoched from 1s before the onset of the first sample to 6s after it, thereby covering the entire trial of 8 samples (including masks). The data were then visually inspected to remove trials containing non-stereotypical artifacts, and to identify electrodes showing electrical artifacts and therefore requiring interpolation. Following this, the data were bandpass filtered between 1 and 40Hz, and re-referenced to the average signal over all electrodes.

An independent component analysis (ICA) was then conducted using EEGLAB, and the resulting ICA components were visually inspected for artifacts, particularly stereotypical artifacts such as blinks and sustained high frequency noise. Trials with artifacts were excluded from all further analysis, leaving an average of 492 (range 378-544) trials per subject, each consisting of 8 overlapping stimulus events (sample onsets).

Unless otherwise stated, we report statistical tests on EEG data averaged across occipital (O1, O2 and Oz) and parietal (P3, Pz, P4 and POz) electrode sites. We chose this approach because previous studies have identified dissociable patterns of activity over occipital and parietal electrodes in discrete-sample integration tasks (Wyart, Nobre, & Summerfield, 2012). To correct for multiple comparisons across time, we used a nonparametric cluster correction technique, implementing a familywise error with an alpha of α=0.05 (Maris & Oostenveld, 2007).

EEG analyses: encoding

We used EEG to investigate how the quantities predicted by the difference and tally models were encoded in neural signals, with a view to arbitrating between them, using a 'model-based' approach to analysis of brain imaging data. Both of these models made identical predictions about choice behaviour, but made different predictions about the neural activity that would accompany each sample. Rather than calculating event-related potentials, we estimated how decision information was encoded in EEG signals using a single-trial approach. This "encoding" methodology involves using parametric predictors (such as DU_{model}) within a general linear regression model to predict the sample-to-sample variability in the EEG signal, at successive timepoints (-100 to 700ms) surrounding the onset of each sample.

First, we took model-predicted quantities $|DU_{difference}|$ and $|DU_{ratio}|$ and standardized these by z-transformation. Using rectified decision updates $|DU_k|$ ensures that we identify neural signals that encode absolute decision information, not those that favour one choice over another (i.e. blue vs. pink), as we were aiming to elucidate the nature of the mechanism rather than the nature of the choice per se. (Consider, as an analogy, random dot kinematogram (RDK) motion discrimination tasks: the neural signal of interest from EEG recordings is one that correlates with the coherence level of the

dots, i.e. decision information independent of direction of motion, rather than one which correlates with the extent to which the information favours leftwards vs. rightwards motion). We then regressed these quantities in a point-by-point fashion against the single-trial EEG activity following each corresponding sample. The resulting parameter estimates (slopes of the best-fitting regression line) provide an estimate of how strongly the EEG signal (at each time point over the course of the sample) varies with these model-predicted quantities. This thus allowed us to assess the difference in neural processing of reliability-weighted and non reliability-weighted information. Although the analysis epochs following each sample are overlapping, we took careful steps to ensure that the correlation between the variables of interest between adjacent samples was minimised. Thus because the decision information provided by each sample is sufficiently uncorrelated, responses to adjacent stimulus events (samples) can be disambiguated, much as they can in parametric event-related functional neuroimaging designs (Josephs, Turner, & Friston, 1997).

Subsequently, we repeated this procedure, including in the same regression the two quantities that are predicted by the difference and tally models respectively: the total number of blue and total number of pink balls (tally model) and the absolute difference in blue/pink balls (difference model). The aim of this analysis was to determine which of the two models was best able to account for the neural activity; including these predictors in the same regression ensured that they competed for unshared variance, allowing us to determine whether neural signals scaled more faithfully with the tally of evidence or the difference of evidence.

In all of these analyses, decision information from the preceding and succeeding samples was included as additional nuisance covariates. This helped ensure that the resulting parameter estimates reflected neural encoding on the current sample, and were not corrupted by decision information from adjacent samples that overlapped in time with the epoch. This step, combined with the fact that the partial correlation between adjacent samples was very low as described, meant that we could be confident that the encoding analysis avoided confounding the influence of future and preceding samples. The resulting parameter estimates (slopes) for each timepoint (-100 ms pre-stimulus to +700 ms post-stimulus) were then averaged across samples and entered into a second-level statistical analysis for comparison at the group level. Regions of time (and space, i.e. electrodes) where these curves deviate reliably from zero across the cohort indicate where decision information is reliably

encoded in neural signals. This is process is also detailed in earlier publications using this method (Wyart, Nobre, & Summerfield, 2012).

EEG analyses: decoding

In a subsequent analysis step, we assessed how the strength of the relationship between decision information and neural signals assessed above (EEG encoding) predicted participants' choices. This analysis step is closely related to the calculation of "choice probabilities" of single-cell recording data (Nienborg & Cumming, 2010) and to an analysis of the psychophysiological interaction (PPI) between behavioural and neural variables in fMRI analysis (Gitelman, Penny, Ashburner, & Friston, 2003). Here and in previous publications, this analysis is called "EEG decoding" (Wyart, Myers, & Summerfield, 2015; Wyart, Nobre, et al., 2012) because it allows us to quantify how fluctuations in EEG *encoding* of DU are 'decoded' in downstream brain structures, and consequently manifest in choice. Such an analysis involves the use of multivariate parametric regression to quantify the extent of the modulatory influence of the EEG signal on the relationship between DU_{model} and choice. A decoding analysis allows us to see whether samples that are encoded with higher than average strength (in other words, with positive residual variance) are more predictive of choice than those which are encoded with lower than average strength; if they are we would see a clear decoding curve (i.e. a significant effect of the residual EEG signal on choice). This method of estimating how the single-trial relationship between input (psychological variable) and brain activity (physiological variable) predicts choices allows a more direct measurement of how brain activity mediates the link between stimulus inputs and the weight or influence that a sample of information wields over choices.

In order to calculate these "decoding" curves, we took the timecourse of the unexplained variance (residuals) of the regression of DU_{tally} and $DU_{difference}$ on the EEG signal. We then entered this quantity, r, into a probit regression, alongside decision information, as a predictor of participants' choices for each sample *k* and timepoint *t*:

$$p(pink) = \phi\left[b + \sum_{k=1}^{8} \omega_k \cdot DU_{k,model} + \sum_{k=1}^{8} \omega_{k,t} \cdot DU_{k,model} \cdot r_{k,t}\right]$$

Once again, averaging across samples and participants afforded a grand average and the opportunity to conduct group-level statistics. Positive deviations of r from zero indicate times at which brain activity not only scaled with decision information, but did so more strongly than on average (i.e. the gain of encoding of a particular sample was above average in comparison to the overall encoding curve). If this stronger neural encoding of the decision update DU results in a stronger effect of the decision update on choice, then this should show as a significant positive interaction between r and DU, as reflected in the weight $w_{k,t}$. The methods used here have been used successfully in previous papers (Wyart, de Gardelle, et al., 2012; Wyart et al., 2015).

EEG: lateralised beta-band activity

Based on previous studies, which have shown that oscillatory activity in the beta-band range accompanies the buildup of information to a decision threshold, we investigated the encoding of momentary and cumulative decision information in time-frequency transformed signals. We measured how signed decision updates (DU_{tally} and $DU_{difference}$) and the corresponding signed decision variables (DVs) were encoded in lateralised beta band activity (~10-40Hz), using a comparable technique to the encoding regressions above, in which these quantities were regressed together against relative lateralised single-trial spectral power over the motor cortex in 10 logarithmically-spaced frequency bands between 9 and 43 Hz. The response made with the left index finger always corresponded to “blue” whilst the right index response always corresponded to “pink”. Thus the use of *signed* decision updates for analysing lateralised beta-band activity is crucial in this analysis because unlike the signals interrogated in previous encoding analyses, here the two alternatives (pink vs. blue) map onto putative neural signals that can be disambiguated at the whole-brain level using scalp EEG (i.e. hemispherically lateralised patterns of beta-band desynchronisation over the motor and premotor cortices). For each subject, we computed the inter-hemispheric difference in beta activity at lateral central electrodes by subtracting the spectral log-power of (CP+CP3) from (C4+CP4).

Control analyses

In a further effort to determine that our results from the main experiments were not being driven by low-level visual properties of the stimuli, we took the number of pink, blue and white balls respectively

on each sample, and standardised these by z-transformation. We then asked whether behavioural choice was significantly influenced by the white dots, and whether neural signals encoded more strongly those quantities that were decision-relevant (pink and blue balls) relative to those that were irrelevant (white balls),

Previous studies that have focussed on the neural representation of number have found a parietal event-related potential that reflects the difference in magnitude (i.e. the total number of stimuli) between one group of dots and the next (Piazza & Izard, 2009; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004). To control for similar effects in our data, we regressed the absolute difference in total number of dots (blue+pink+white) between one sample and the next against the EEG signal (using the regression methods described above, including samples 2-8 in the stream). In particular, we were interested in whether any decision-related signals observed in our experiment could be trivially explained by previously described neural adaptation to number.

Behavioural control experiment

In the experiments described above, all dots were equally sized, and so the number of dots was correlated with the pixel area of the coloured dots on the display screen. Thus one question that arises is whether choices are driven principally by low-level visual properties of the stimuli, i.e. the area (number of pixels) that the coloured dots take up on screen, or by the number of dots *per se*. In order to arbitrate among these possibilities, we conducted a further behavioural experiment in which we varied the size of the pink and blue dots from sample to sample. This ensured that the number of dots and the number of pixels favouring each choice were decorrelated.

The design (including timing of stimuli and number of trials) of this behavioural control task, which was conducted on a new cohort of 19 participants, was identical to the first EEG experiment except for the size and position of the dot array. Dots were randomly spatially distributed (minimum separation 10 pixels) among 46 possible locations within a circular aperture of 500 pixels diameter. The diameter of the blue dots on each sample was randomly selected from a range of 11-45 pixels (35 possible sizes), and the same process was applied independently to both the pink dots and the white dots on each sample, thus ensuring that the pixel area of each colour separately on each sample, as

well as the total filled pixel area, was decorrelated from the respective number of dots (Pearson's correlation ~0.65).

This control allowed us to regress both the number of dots and the size of the dots on each sample against choice, thereby allowing us to determine whether sample size exerted an influence over and above the physical size of the dots on screen.

2.3 Results

Behaviour

Participants chose the correct urn on 83±6% of trials, with reaction times averaging 407ms. After the addition of performance-limiting Gaussian noise to the DV, all three models achieved comparable accuracy to humans (85% for the tally and difference models, and 83% for the ratio model) and were able to predict psychometric functions well. We calculated how participants' choices varied as a function of the z-transformed decision variables predicted by the tally and difference models and the ratio model, $DV_{difference}$, DV_{tally} and $DV_{ratio.}$ The models performed equivalently (**Figure 2.2a**)**.** The tally and difference models behaved in line with the optimal solution, and both (as described) made equivalent predictions with the exception of noise.

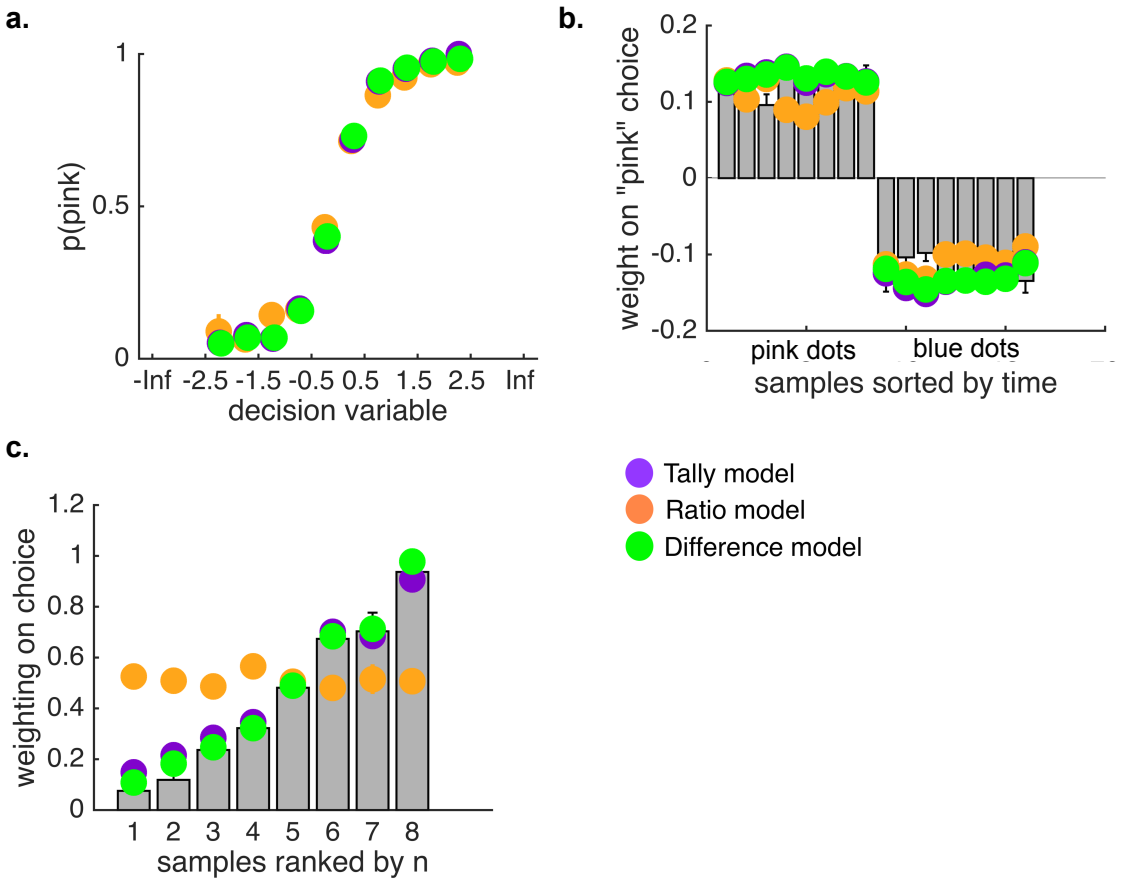

Figure 2.2. Behavioural results and model predictions. (A). Probability of selecting "pink" as predicted by the DVs for the ratio model (orange circles) and the two reliability-weighted models (tally model, purple circles; difference model, green circles); a DV below zero corresponds to responding "blue" and above zero corresponds to a "pink" response. (B). Impact (beta coefficient) of each sample on "pink" choices, ranked by the serial position (1-8) as a function of the number of pink balls (first eight bars) and the number of blue balls (last eight bars) in the sample. Estimates were generated using probit regression. Grey bars show human data and orange-/purple-/green- coloured circles show predictions of the ratio/tally/difference models respectively. (C). Weights (beta coefficients) given to each sample as ranked by sample size (smallest to largest) in evidence integration, calculated using probit regression. Grey bars correspond to human behaviour, coloured circles show model estimates for weight given to the ranked samples. The ratio model, given that it does not reliability-weight the samples, is flat across ranks. In all figures, bars show SEM.

We next used probit regression on the behavioural data to estimate the impact that the number of pink and blue balls in each sample (1-8) had on choice, as a function of its serial position. As expected, "pink" choices were predicted positively by the number of pink balls, and negatively by the number of blue balls (**figure 2.2b**); this analysis suggested that all the samples contributed at least in part to the decision (all grey bars deviated significantly from 0).

Finally, in order to test our main hypothesis, we determined whether or not subjects accounted for sample size (number of dots) when making decisions. To this end, we used probit regression to estimate the impact that each sample had on choice, ranked not by its position, but by the total number of coloured dots, n_k (= $d1_k$ + $d2_k$), i.e. by its overall reliability. We used DU_{ratio}, the decision update as calculated by the ratio model, as a predictor. This ensured that the resulting coefficients for observers who did not weight information by sample size would be flat over ranks of n_k, whereas an observer who weighted information by reliability would show a profile of steadily increasing weights (note that a mathematically equivalent alternative would be to use the statistically optimal solution as a predictor, in which case an ideal observer would show a flat profile of weights, whereas those for an observer who integrated the ratio of evidence would decline with n). Consequently, for the ratio model (orange points, **figure 2.2c**), the weights did not vary with ranks of n_k, but the weight given to each sample ranked by n_k grew steadily for both the models that arrived at the statistically optimal solution (tally model, purple points, and difference model, green points). Again, the values predicted by these latter two models are equivalent to the statistically optimal solution, with any residual variability due to the noise term σ.

Critically, the impact that each sample wielded over choices for humans depended on the sample size, just as it did for the statistically optimal solution and the two models that approximated it; this was confirmed by an ANOVA on the weights over ranks ($F_{(7,240)} = 51.6$, $p < 0.001$). The beta weight in the human data for the largest samples was significantly higher than that of the smallest samples: $t_{(30)} = 13.1$, $p < 0.001$. In the subsequent neural analyses, we seek to distinguish which of the two approximations of the optimal (binomial) solution best describes human brain activity during performance of the task. The reasoning behind this approach is that it is implausible that neurons explicitly compute binomial probabilities, but rather derive the solution via a computationally tractable mechanism, such as those described by both the tally and the difference models. Furthermore, it allows us to tease apart the difference and tally models that make indistinguishable behavioural predictions.

EEG: encoding of decision update (DU) in broadband occipital and parietal signals

How did humans achieve near-optimal performance on the task? Although the tally and difference models both derive the statistically optimal solution with predictions that match human behaviour, they make different predictions about the quantities that are computed *en route* to a decision. We thus analysed the link between decision information, neural signals and choices to provide an insight into the mechanisms by which humans were making decisions.

First, we sought to correlate EEG activity with the decision update DU (i.e. the momentary information conveyed by each sample) predicted by each of the models. Initially, we focussed on comparing the ratio model to the difference model, only the latter of which predicts reliability-weighted behaviour. To this end, we regressed $|DU_{difference}|$ and $|DU_{ratio}|$ separately against the single-trial EEG data for each sample, and averaged over the resulting coefficients at each post-stimulus timepoint. The use of rectified predictors allows us to identify neural signals that correlate with the absolute decision information, rather than favouring one choice over another (e.g. pink over blue). Standardization ensured that the resulting coefficients were directly comparable. Consistent with participants' behavioural tendency to act as if they were weighting the evidence by its reliability (**Figure 2.2**) the response to $|DU_{difference}|$ was significantly greater than that to $|DU_{ratio}|$ (**Figure 2.3a).** In other words, behavioural data indicate humans pursue a near-optimal strategy, and neural data suggest that they do so by integrating reliability-weighted evidence rather than just the ratio of evidence on each sample.

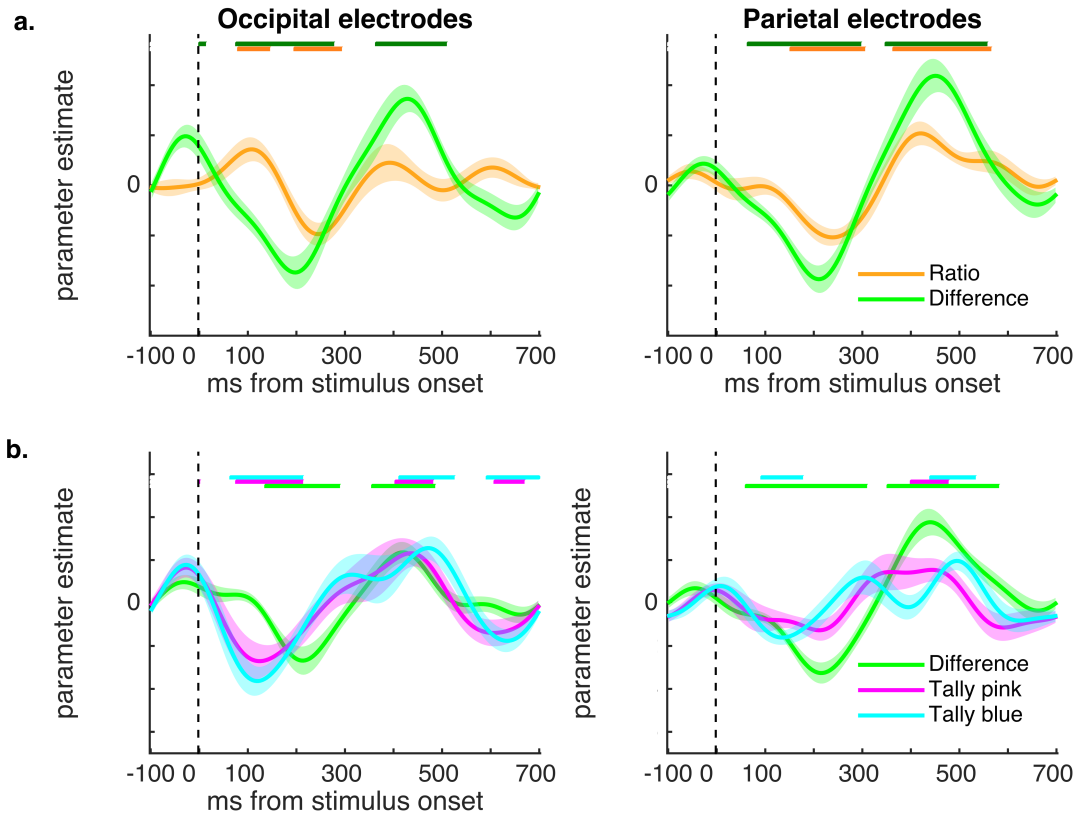

Figure 2.3. Neural encoding curves. (A). Curves showing correlation between the predictions of the ratio model (orange curve) and the difference model (green curve), which approximates the binomial solution, with the EEG data after each sample. Shaded areas in the curves denote *SEM.* The bars at the top show periods from stimulus onset in which the curves consistently deviate significantly from zero across participants for difference (green) and ratio (orange) models. (B). Correlation between EEG signal and the DUs predicted by the tally model (blue and pink ball totals, blue and pink curves) and the difference model (blue and pink balls, green curve). Shaded areas in the curves denote *SEM.* The bars at the top show periods from stimulus onset in which the curves consistently deviate significantly from zero across participants for difference (green bar) and tally (blue and pink bars) models. Note that, for occipital electrodes, there is a period of significant activation corresponding to the tally model after stimulus onset, followed by significant activation corresponding to the difference model around 200ms, as though all evidence is being processed initially followed by the formation of the decision-relevant evidence signal. For all panels, statistics were computed using a nonparametric cluster-correction technique implementing a familywise error correction with an alpha of 0.05 (see methods).

Subsequently, we asked whether humans performed near-optimally by (i) adding up the differences in numbers of balls and deciding whether this quantity was greater or less than zero (difference model), or (ii) adding up the total information in each stream in parallel, and deciding which was greater (tally model). The former account predicts that participants compute $DU_{difference}$ on each

sample; the latter account predicts that only DU_{tally} is updated on a sample-by-sample basis. We thus entered $DU1_{tally}$, $DU2_{tally}$ and $|DU_{difference}|$ as competitive predictors of the EEG signal at each scalp electrode and timepoint from 100 ms preceding sample onset to 700 ms following its onset (see Materials and Methods). We observed that the neural variance accounted for by $|DU_{difference}|$ in parietal electrodes outweighed that accounted for by $DU1_{tally}$ and $DU2_{tally}$, in that the beta weights were larger for the former and reached significance for longer, with a first negative deflection peaking at 230 ms and a second positive deflection at 470 ms after stimulus onset (**figure 2.3b,** right panel). It is important to note that these curves are not event-related potentials, but single-trial estimates of the encoding of decision information by brain activity. Interestingly, the same regression in occipital electrodes (**figure 2.3b,** left panel) showed an initial negative deflection for $DU1_{tally}$ and $DU2_{tally}$ followed 150 ms later by a negative deflection for $|DU_{difference}|$. One might expect, even under the difference model, that there would be an early representation of the total number of pink and blue dots, as this quantity is required in order to calculate the difference. Thus, our interpretation of this finding is that the initial encoding of the total information is followed by the emergence of the decision-relevant quantity $|DU_{difference}|$.

Repeating the above analysis at each electrode provided a topography of encoding of $|DU_{difference}|$ and DU_{tally} across the scalp, which we collapsed into bins of 100ms (**figure 2.4,** note that the two quantities $DU1_{tally}$ and $DU2_{tally}$ are combined for ease of presentation given their similarity). The spatial distribution of the resulting weights can be interpreted as the spatial topography of the underlying decision-related component in the EEG signal. The resulting patterns of significant activation were in accordance with the results from the regressions above, with the response to $|DU_{difference}|$ (bottom row) outweighing that to DU_{tally} (top row). In other words, these neural observations suggest that participants use a strategy that involves encoding the difference in information provided by each sample. This allows them to derive the near-optimal solution, weighting information by its reliability.

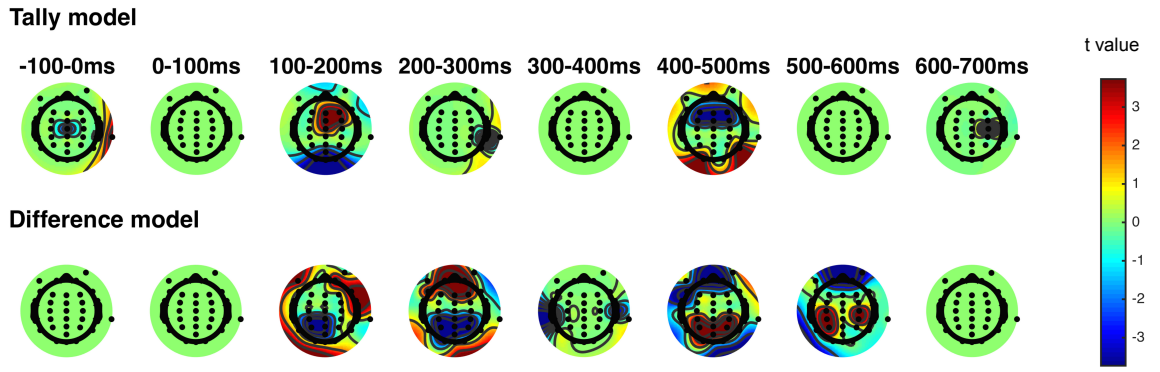

Figure 2.4. Scalp topographies. Scalp-wide significant correlations between the EEG signal predicted by the tally model (top row) and the difference model (bottom row). Note that the tally model predicts that two quantities are encoded (number of pink balls and number of blue balls); these two quantities have been combined in this figure for ease of viewing but plotting them separately yields a qualitatively similar pattern. The plots show the t values corresponding to times and regions at which (on average in that time bin) the correlation for each model deviated significantly from zero.

EEG: decoding of decision update (DU) in broadband parietal signals

The encoding regressions suggested that more of the neural variance was accounted for by the difference model, which encoded the relative evidence on each sample, than by the tally model, which encoded the absolute evidence. To support this notion further, *decoding* analyses – which are able to link the psychological variable (human choice) with the physiological variable (neural encoding of DU in the run up to choice) – were used to assess the relationship between the residual variance not accounted for by the encoding regressions, and human choice (see Materials and Methods). This technique allows the link between the neural transformation of DU and choice behavior ('choice probabilities') to be made more explicit. Decoding analysis asks how residual variance in the encoding (described above) of model-predicted quantities impacts choices, over and above the influence of stimulus choices. Whilst the decoding analysis for the tally model showed no significant effect (**figure 2.5a**), the time courses of the decoding regressions for $DU_{difference}$ were in accordance with those seen in encoding in parietal electrodes (**figure 2.5b**). This analysis therefore provides support for the difference model over the tally model, and suggests that variation in encoding of the difference signals is predictive of participants' responses. We note that because it shows the link between neural activity and behaviour, this finding also rules out spurious explanations for the observed decision encoding

curves, such as the view that the apparent relationship between decision information and brain activity is somehow secondary to differences in attention or arousal.

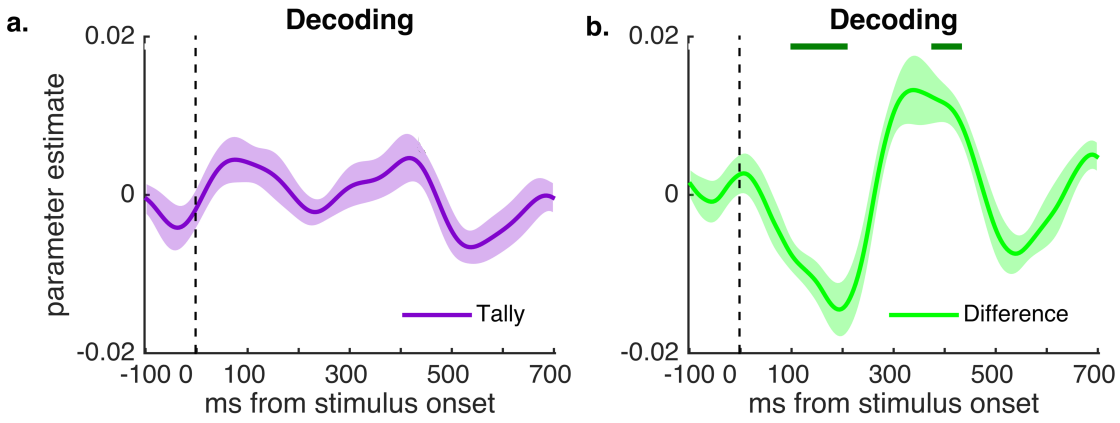

Figure 2.5. Neural decoding curves. The unexplained variance (i.e. the residual error) in the EEG signal from the encoding regressions in parietal regions for the tally model (A, purple curve) and the difference model (B, green curve) were used as predictors of choice in the decoding regressions, allowing us to link neural activity (in parietal electrodes) with behaviour choice. See methods for more details on this procedure. Shaded areas in the curves denote *SEM.* The dark green bar shows periods in which the correlation significantly deviated from zero in the parietal signal; note that, for the tally model, there was no such period of significance.

EEG: lateralised beta-band activity

Previous studies have observed that oscillatory activity in the beta-band range accompanies the build-up of information to a decision threshold (O'Connell, Dockree, & Kelly, 2012). Next, thus, we measured how lateralized beta-band activity over the motor cortex varied with the cumulative decision evidence in favour of either choice. To this end, we wavelet-transformed EEG data into its spectral components in 10 logarithmically-spaced frequency bands between 9 and 43Hz (i.e. encompassing the approximate beta band range extending into lower gamma). First, we confirmed that lateralized beta band activity was present in the preparation of the motor response (made from 3500ms after the onset of the first stimulus in each trial) by computing the inter-hemispheric difference in EEG activity in the 9-43 Hz range between the lateral central electrodes (see Materials and Methods). The results confirmed that the difference in power in preparation for responding “pink” (the choice made with the right index finger) minus responding “blue” (made with the left index finger), was positive in the

contralateral hemisphere and negative in the ipsilateral hemisphere (**figure 2.6a)**, with a focus between ~20 and ~30 Hz.

Finally, we tested whether beta-band activity correlated with the accumulated decision-relevant evidence predicted by each of the models. To this end, we conducted further regression analyses in which the momentary ($DU_{difference}$) and cumulative ($DV_{difference}$) predictions, as well as the momentary (DU_{tally}) and cumulative (DV_{tally}) predictions, were both entered as predictors of the inter-hemispheric difference in lateralized log-power in the 10 selected frequency bands (see Materials and Methods). Note that for these analyses we used the signed decision quantities (not rectified) in order to predict the leftwards vs. rightwards response. Consistent with the view that participants solve the task by integrating a difference signal, the cumulative difference signal $DV_{difference,}$ but not the cumulative tally signal DV_{tally} was a reliable predictor of beta-band lateralization over the motor cortex, as the relationship between the lateralised beta band activity was significantly associated with $DV_{difference}$ from 100 ms onwards (**figure 2.6b** and **c,** rightmost panels), but was not significantly associated with DV_{tally} (**figure 2.6b** and **c,** leftmost panels).

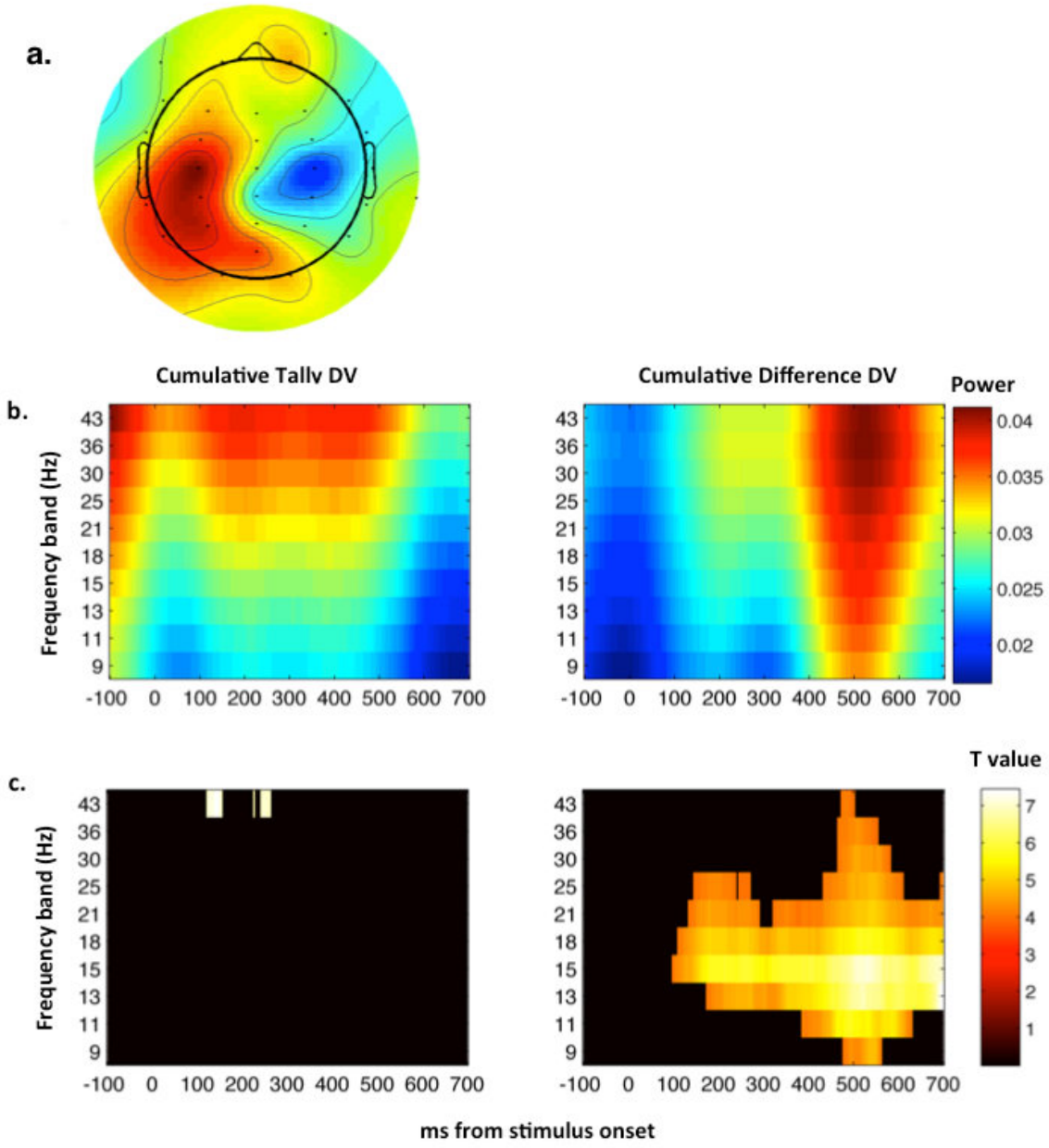

Figure 2.6. Time–frequency analyses. (A): Interhemispheric difference in log spectral power at 9–43 Hz (i.e., mainly overlapping the beta band frequency, approx. 15–30 Hz) between "blue" choices, which were always made with the right index finger, and "pink" choices, made with the left index finger. The resulting quantity was positive in the contralateral hemisphere to the hand with which the choice was made and negative in the ipsilateral hemisphere. The period displayed is 2.6–3.2 sec after first stimulus onset, that is, close to execution of motor response. (B): Tally and difference model DVs were generated by computing the cumulative sum of the model-derived sample- by-sample quantities DU_{tally} and $DU_{difference}$, respectively, across the trial. Note that the tally model predicts that two quantities are encoded (number of pink balls and number of blue balls); these two quantities have

been combined in this figure for ease of viewing. The resulting predictors were then regressed against the log spectral power difference between centroparietal electrodes C3 + CP3 and C4 + CP4, that is, a motor region that should correspond with motor preparatory activity. The left panel shows the correlation between the predictions of the cumulative tally DV and the difference in power between the two hemispheres, and the right panel depicts the same but in relation to the difference model DV. (C): The left and right panels show times/frequencies at which the respective quantities plotted in B (the middle left and right panels) deviate significantly from zero.

Control analyses

Next, in order to ensure further that the neural responses were driven by decision-relevant information rather than other, low-level factors, we first measured how both behaviour and brain activity correlated with the number of white distractor dots in comparison to pink and blue dots. These analyses allowed us to distinguish the behavioural and neural response to decision information (blue and pink dots) from that elicited by task-irrelevant sensory input (distractor white dots). Entering these quantities (white, blue and pink dots) together into the regression ensured that they competed for unshared variance. In the behaviour data, the response to white dots was minimal in comparison to that evoked by reliability-weighted coloured dots (**figure 2.7a**). In the neural data, the time course of the resulting regression coefficients showed that the number of white (distractor) dots drove a visual response (**figure 2.7c**) peaking around 250ms post-stimulus. Whilst the neural encoding of the white dots was significant, it followed a qualitatively different pattern from the encoding responses to the decision-relevant quantities (pink and blue dots; indicated in **figure 2.7c** by pink and blue curves respectively) in both parietal and occipital electrodes. Combined with the behaviour analysis showing limited effect of white dots on choice, the qualitatively different neural encoding pattern suggests that the white dots were being processed via a different, decision-irrelevant mechanism. The responses to the coloured dots were characterized by negative and positive deflections peaking at ~100 and ~400 ms post-stimulus respectively.

We also examined the relationship between brain activity and the difference in magnitude (i.e. the total number of dots on each sample, both decision relevant (blue and pink) and decision irrelevant (white)) from one sample to the next, regressing the absolute difference in magnitude from sample to sample against the EEG signal over parietal electrodes (see Materials and Methods). Perhaps surprisingly, we found no correlation between the change in magnitude and parietal EEG signals, with

no time window reaching statistical significance (**figure 2.7d**). This suggests that our neural data are dominated by decision-related effects rather than the influence of passive adaptation to number.

Control experiment

Finally, in order to determine whether the observed effects of sample size were driven by the number of dots, rather than low-level visual signals such as the total pixel area taken up by the coloured dots, we conducted a further control experiment. In this experiment, a new cohort of participants (n=19) performed the same urn choice task, with one difference: the pink and blue dots were different sizes, ensuring that the pixel area taken up by the dots and the number of dots were dissociable on each sample, for each colour individually (correlation between number of pink dots and pink pixel area: r= 0.65, correlation between number of blue dots and blue pixel area r= 0.66) and overall without white dots (r= 0.68) and with white dots (r=0.65). This enabled us to investigate sample size effects using the same regression approach, in which we predicted choices as a function of the ratio of dots, sorting samples into predictors by their sample size (see above) but now partialling out that portion of the variance that could be ascribed to pixel area. To achieve this, we entered the sample-wise pixel area for blue and pink categories as additional regressors into the design matrix. The resulting coefficients are shown in **figure 2.7b**. The results show that whilst decisions are partly influenced by number of pixels, the effect of larger samples having greater influence on choices remains robust to the inclusion of these additional regressors, following the same pattern of increasing weight on choice with increasing sample size that was seen in the main analyses. This was confirmed by an ANOVA on the weights over ranks: for the weights corresponding to sample size, $F_{(7,144)} = 32.6$, $p < 0.001$. As before, the beta weight in the human data for the largest samples was significantly higher than that of the smallest samples: $t_{(18)} = 7.8$, $p < 0.001$. However, the same was not the case for the beta weights corresponding to pixel area: although the weights deviated significantly from zero, the weights did not deviate significantly from each other (F=0.71, p=0.66 for blue pixel area and F=0.67, p=0.7 for pink pixel area).

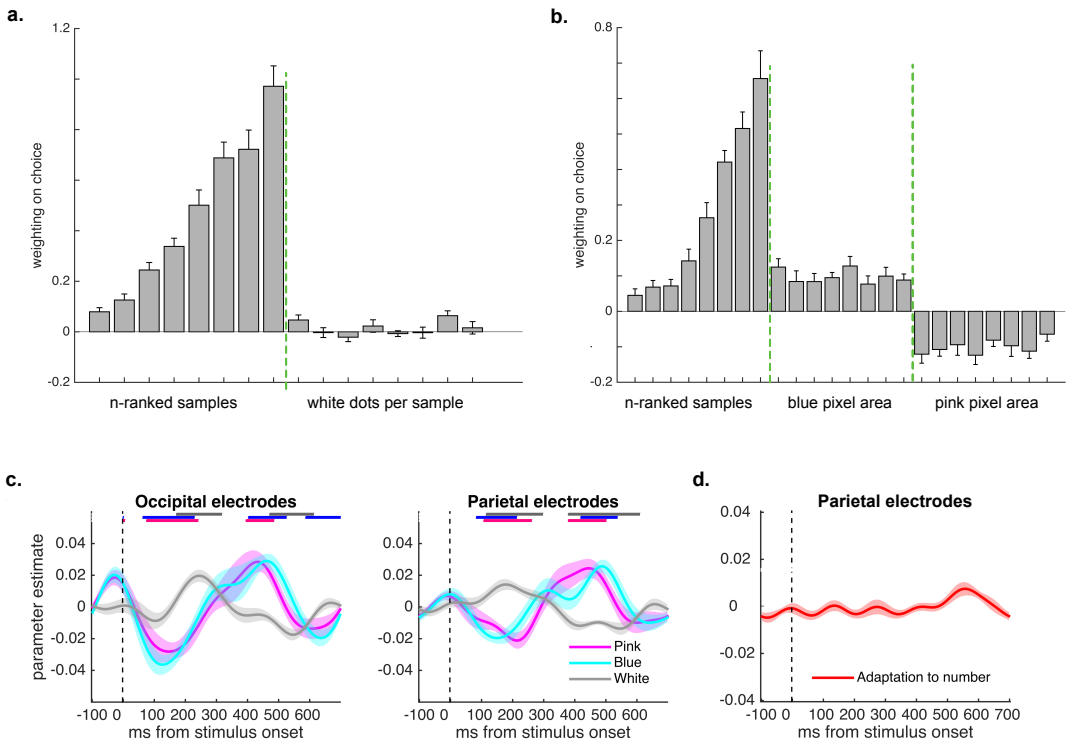

Figure 2.7. Control analyses. (A): the first eight grey bars show weights (beta coefficients) given to the coloured (decision-relevant) dots in each sample, ranked by sample size (smallest to largest). The second set of eight grey bars show the weight given to white (decision-irrelevant) dots in each sample. The response to the coloured dots clearly outweighs the response to the white dots. (B): grey bars depict the weight (beta coefficients) on choosing "blue" of various aspects of the decision array. The first set of eight bars show weights given to each sample as ranked by sample size (smallest to largest), the second set of eight bars show the weight of total pixel area of blue dots per sample, and the final set of eight bars show the weight of the total pixel area of pink dots per sample, on choosing "blue." All quantities were entered into the same (probit) regression to ensure they competed for unshared variance. In all, black bars denote SEM. (C): correlation between decision-relevant (blue and pink dots) and decision-irrelevant (white dots) information and the EEG signal in occipital electrodes (left) and parietal electrodes (right). Shaded areas in the curves denote SEM. The bars at the top show periods in which the curves consistently deviate significantly from zero across participants and trials; the colours of these bars correspond to the colours of the curves they reflect. (D): we regressed a quantity corresponding to the difference in number of dots (blue + pink + white) between one sample and the next against the EEG signal over parietal electrodes. The curve shows that the correlation between this quantity and the neural activity was small with no time points reaching significance. Shaded areas in the curve denote SEM. This suggests that the parietal encoding signal (see Figure 2.3) is indeed likely to reflect decision-related effects, rather than the influence of passive adaptation to number.

2.4 Discussion

Humans making decisions are often faced with information that is indistinct, weak, or ambiguous. During sensorimotor choices, humans rely more heavily on inputs coming from sources that are clearly discernible, for example by weighting haptic information over visual information when the latter has been corrupted by noise (Ernst & Banks, 2002). In one sense, it is remarkable that human behaviour has been optimised over evolutionary history to weight information by its reliability, just as a statistically ideal observer should. On the other hand, it is perhaps natural that human decisions rely more on information that can be easily detected than that which cannot. Even the simplest tasks can contain more than one source of uncertainty. Let's consider another courtroom analogy. To a juror evaluating evidence, a witness whose testimony is rendered incomprehensible will be unlikely to hold much sway, compared to one that can be clearly understood (uncertainty due to sensory noise). However, even where testimony is clearly perceptible, a juror might question whether a lone witness (rather than, say, ten consistent testimonies) is sufficient to condemn the accused (uncertainty due to knowledge of sample size). Here, we tested whether humans integrate information optimally when uncertainty is particularly influenced by the size of the sample. Unlike previous studies using description-based scenarios, we found that participants automatically weighed the information by sample size, behaving in a fashion that resembled an ideal observer.

Previous research has revealed that humans often disregard the quantity or quality of information when making judgments (Griffin & Tversky, 1992). This 'sample size neglect' can be considered as a special case of a more general bias by which humans neglect the duration or extent over which information is available (Tversky & Kahneman, 1974a). However, this classic research presented decision problems in the form of descriptive scenarios, making it unclear whether optimality was simply limited by the format of decision information (description- vs. experience-based). Here, we assessed human decisions using an expanded judgment task in which discrete samples of information arrived in sequence, following which participants made a categorical judgment about their provenance. We found that participants paid more heed to samples that offered better-quality information: for example, decisions were more influenced by samples on which there were 12 balls, rather than 3 balls, even if the ratio of blue to pink balls was 2:1 in both cases. In other words, this classic demonstration of suboptimal behaviour from the decision-making literature may indeed arise

simply due to the format in which information was presented. This has important implications for a range of real-word situations, including medicine, economics and public policy, where sample size neglect might lead to poor or biased decisions.

The optimal solution to the urn-and-balls problem described here is to estimate the binomial odds ratio that the balls were drawn from either of the two urns, and decide according to the sign of its logarithm. Note that samples were presented at 3 Hz, too fast to rely on explicit mathematical calculation; any integration must be of approximate number, or magnitude (Piazza & Izard, 2009). Thus, although humans behaved near-optimally, it is implausible that they were explicitly computing binomial probabilities to solve the task. Rather, there are two mechanisms that arrive at the binomial solution that participants could have used. The first is simply to tally up the approximate number of pink and blue dots independently, and respond whichever is greater (tally model). The second is to estimate the approximate difference between pink and blue dots on each sample, and respond according to whether it is greater or less than zero (difference model). Where the experimenter, rather than the observer, determines the viewing time that precedes choices (as in this experiment), these models make very similar predictions (identical, except for added stochasticity) regarding behavioural data, but differ in their neural predictions. The computational approaches described by the tally and difference models are related respectively to the 'race' and 'diffusion' accounts of the integration process that is a prelude to human categorical choices (Ratcliff & McKoon, 2008; Vickers, 1979).

The tally model requires that independent totals of blue and pink dots are registered, but not their relative difference. Nevertheless, we witnessed neural signals over the parietal cortex that scaled with the difference in the total number of blue and pink dots, even when other confounding factors had been taken into account. This supports the difference model but not the tally model. Indeed, a closer look at the encoding of information over occipital electrodes suggested that the number of blue and pink dots was encoded early (100ms), followed by the difference signal, exactly as if the brain first estimated independent totals for each sample, and then compared them. This is reminiscent of the successive encoding of absolute and relative economic value in magnetoencephalographic activity observed during a gambling task (Hunt et al., 2012). In the parietal cortex, the decision information predicted by the difference model exhibited a temporal profile characterized by an early dip (at ~220 ms post-stimulus) and a later peak (at ~450 ms). The same pattern of EEG activity has previously been

shown to scale with decision information in a task involving discrimination of the mean orientation of a stream of tilted gratings, with a negative deflection at ~250ms and a positive peak at ~500ms following each sample (Wyart, Nobre, et al., 2012).

A wealth of research in humans and non-human primates has implicated the parietal cortex in the representation of approximate number (Piazza & Izard, 2009). At the single-cell level, individual neurons show bell-shaped tuning curves over the number line (Nieder & Miller, 2003). In human imaging studies, neural signals localized to the parietal cortex scale with the numerical disparity between two successively- or simultaneously-occurring stimuli (i.e. sets of dots), even in the absence of an overt estimation task (Piazza et al., 2004). The focus of our experiment was not to distinguish accounts based on number from those based on magnitude per se. Nevertheless, it is unlikely that our neural findings simply reflect low-level adaptation to number or magnitude, for four reasons. Firstly, a control experiment in which the pixel area taken up by the coloured dots and the number of coloured dots were decorrelated still showed a strong sample size effect, even when the lower level factor (pixel area) was included as a nuisance covariate in the regression. Secondly, the gain of encoding of number was qualitatively different for blue and pink stimuli, which were decision-relevant, than for white stimuli, which were not. Thirdly, our decoding analysis suggested that the strength with which neural signals encoded the difference in number of pink and blue dots was predictive of later choices, as would be expected of a decision signal. Fourth, we observed a distinct neural signal over motor regions that reflected the build-up in magnitude differences over time. These findings suggest that the parietal signal observed here instead reflects the relative difference between decision information in favour of either category. One way of linking our findings with this literature is to assume that in the absence of an overt task, participants implicitly compare information arriving in sequence, calculating their difference as an implicit decision signal. Indeed, a similar pattern of adaptation to numerosity is obtained when participants make same-different judgments on successive stimuli. In other words, the representation of numerosity in the parietal cortex might reflect a more general representation of the magnitude of the information relevant for a decision (Wyart, Nobre, et al., 2012).

In sum, we show that in a simple perceptual decision-making task, humans integrate approximate number in a near-optimal (reliability-weighted) fashion. Our study has implications for research in a number of domains, including economics and medical diagnosis, in which it is widely assumed that

humans estimate numerical quantities in a biased and suboptimal fashion. Humans can make near-optimal judgments about values or other numerical estimates if the information is presented in an appropriate format.

3. Selective integration of information in multi-feature decision making

In the previous chapter, we saw that in a simple categorisation task, humans made decisions that conformed to the statistically optimal solution, integrating all available evidence and employing it according to its reliability. However, whilst this task involved the need to integrate evidence across a series of eight samples, and as such featured 'late' integration noise – noise that corrupts the choice at stages that occur after the sensory level, for example during evidence integration (see chapter 1) – the paradigm used was a simple 'single-feature' task, i.e. one that has only one type of evidence in favour of each option, so the degree of late noise might be limited. In tasks in which late noise is more prevalent, the strategy that leads to the greatest chance of a correct decision involves a "rationally suboptimal" strategy: the selective integration of evidence with a momentarily higher value (Tsetsos et al., 2012; Tsetsos et al., 2016). Here, we investigated whether a selective integration process may underpin decisions between categories that are characterised by more than one feature (a 'multi-feature' task). Using a second variation of the classic urn and ball paradigm employed in the previous experiment, participants again indicated whether a series of eight samples of coloured balls were drawn from an urn that had a 60:40 ratio of category A:B or vice versa, but this time the categories themselves comprised more than one feature (colour). In this more complex task, we show that performance is best captured by a selective integration model that exhibits selectivity for the locally winning features within a category, but not between categories. The use of concurrent EEG recording revealed the first direct neural evidence for this type of selective integration, again at the subcategory but not the category level. Whilst selective integration at the category level has been shown to lead to a higher chance of being correct when late noise is high, this does not hold for the subcategory selectivity shown in this task. This suggests that whilst selective integration may be a good general strategy for maximising accuracy, it does not improve decisions in all circumstances.

3.1 Introduction

Good decisions require the lossless accumulation of evidence over time (Bogacz et al., 2006; Wald & Wolfowitz, 1949). A canonical view in the field holds that humans integrate evidence in line with a Bayesian (optimal) observer, limited only by decision noise arising from ambiguity in stimulus encoding and transduction (see chapter 1). However, agents are also prone to down-weighting or ignoring part of the available information, leading to apparently suboptimal behaviour. For example, in a multi-sample categorisation task that involved choosing between two streams of items, individuals consistently prioritised information that had locally higher attribute values (e.g. the larger number when judging which of two streams of digits had the higher average). This strategy, known as 'selective integration', leads to transitivity violations (Tsetsos et al., 2016) and other preference reversals (Tsetsos et al., 2012).

Despite leading to violations of rational choice strategy, the authors of these studies showed that selective integration can be optimal (i.e. reward-maximising). To understand this seemingly paradoxical statement, it is important to consider the different stages at which decisions may be corrupted by noise. Noise can be external (i.e. in the stimulus itself) or internal (i.e. physiological noise), with dissociable impacts on choice (Wyart, Nobre, et al., 2012). Furthermore, internal noise can occur early, in sensory processing, or later, at the stage of evidence integration or choice execution. As shown by Tsetsos (2016), when decisions are corrupted by "late" noise, the best policy may be one that prioritises some information at the expense of other information. Thus, decisions made on the basis of down-weighted information can be "robust" in the face of noise, leading to the best (most likely to be accurate) choices in the given context. This is important, because "late" noise may predominate for many cognitive tasks (Drugowitsch et al., 2016; Herce Castañon et al., 2018).

The selective integration model provides a process-level account that captures this accumulation-with-loss process in a way that leads to robust decisions in the face of late noise. Consider a task in which observers compare two stimuli that are each composed of multiple attributes, by comparing each attribute value in turn. According to the model, after initial sensory processing, the inputs (e.g. an attribute for each of the 2 competing stimuli) are fed through a bottleneck (representing processing constraints) that allows the 'local winner' on each sample – i.e. the attribute with the highest value –

to be processed with enhanced gain relative to the alternative. The information is then accumulated, where it is subject to further 'late' noise (Tsetsos et al., 2016). The selective integration model has behavioural support from categorisation tasks involving integration of symbolic numbers, as well as those involving perceptual features such as bar height, hinting at its ability to account for general performance rather than tasks limited to a specific stimulus. Because the selective integration model postulates the existence of a bottleneck that leads to the incomplete processing of some information, it implies that the selective integration of information reduces neural processing load. However, although the model has behavioural support, this specific neural prediction of the model has not been tested directly.

Furthermore, as of yet it is unclear whether this model is able to generalise to explain more complex multi-feature categorisation problems (i.e. those in which each category is comprised of several subcategories). For example, when choosing between which of two rescue dogs to adopt – a Pomeranian or a Labrador – based on their general appearance, the selective integration model would predict that information about the more attractive dog should be integrated preferentially. However, the 'appearance' attribute may itself be defined by several sub-features: for example, attractive dogs are ones with soft looking fur as well as those who often have an amusingly guilty facial expression. It is unclear whether this selective integration model would extend to predict that winning features *within* an attribute are also preferentially integrated alongside the winning attribute itself. It could be that the extreme fluffiness of the Pomeranian (currently number one on 'The top 12 fluffiest dog breeds that ever floofed'[1]), and the tendency of the Labrador towards a resting guilt face, would be prioritised in the decision process, irrespective of which dog has the better *overall* appearance. Similarly, when comparing personalities, we may also believe that the most loveable dogs are the less intelligent ones, and prefer boisterous animals with booming bark rather than a fussy yap, so we also want to compare these individual features (intelligence and barking style) within the 'personality' attribute to drive our choices. Whilst healthy individuals have been shown to use an 'attribute based' decision strategy – comparing each attribute of the dogs in turn, i.e. comparing their appearance, then their personality – before an 'option based' strategy, which involves comparing the overall qualities of each of the dogs (Farashahi, Rowe, Aslami, Lee, & Soltani, 2017; Fellows, 2006), it is less clear how features *within* attributes are compared when making decisions. The selective integration

[1] Retrieved from https://barkpost.com/fluffy-dog-breeds/

model has not so far been applied to tasks that have involved this additional level of complexity. We may expect that in a multi-feature categorisation task, individuals would selectively integrate winning *features (subcategories)* as well as the overall winning *attribute (category)*, but this remains to be seen.

This leaves us with two unanswered questions: firstly, do we see behavioural signatures of selective integration in multi-feature categorisation tasks, and secondly, do we see neural indicators of selective integration? Here, thus, we used a multi-attribute binary categorisation task in conjunction with EEG recording to address these questions. Participants viewed a series of eight samples of coloured balls drawn from one of two virtual urns, and indicated whether the balls were drawn from the urn with a 60:40 ratio of category A:B, or from the other urn (ratio vice versa). Importantly, the two categories each comprised two subcategories (i.e. 'features', manifested as two different colours). Therefore, each sample had a locally winning category, as well as two locally winning subcategories (one per category). To foreshadow our results, rather than seeing selective integration of the locally winning category (as was seen in the previous, single-feature experiments), we instead saw within-category selective integration of the locally winning subcategory. This was captured by a modified version of the selective integration model that incorporated local down-weighting of subcategory (rather than category) information, and which explained human behaviour and EEG data well.

3.2 Method

Participants

The investigation consisted of two parts: part 1 was a solely behaviour experiment, and part 2 was a replication of part 1 with concurrent EEG recording. Forty-one (solely behaviour) and 20 (EEG) participants, all of whom reported normal or corrected-to-normal vision, were recruited from University College London, the University of Oxford and the surrounding areas (total n=61). All participants provided written informed consent in line with local ethical guidelines. As compensation for their time, participants received £10 (solely behaviour) or £20 (EEG), plus up to an additional £9 performance-based bonus (EEG version only). Data from participants who scored lower than 65% were excluded (n=4 and n=1 from the solely behaviour and the EEG versions respectively). Furthermore, data from the EEG experiment in which there was excessive movement and/or electrical interference in the signal were removed (n=4). This left n=37 for solely behaviour version, and n=15 for EEG version (total n=52). Unless otherwise stated, all behavioural results are collapsed across the two versions.

Task design and stimuli

Participants completed a probabilistic decision making task (**Figure 3.1**) that was a variation of a classic urn and ball paradigm. On each trial, participants saw a series of eight samples of coloured balls (orange, blue, green and pink) drawn from a virtual "urn". The urn was pseudo-randomly selected on each trial, and either contained a 60% predominance of category A (itself comprised of two of the four colours) and 40% category B (comprised of the other two colours), or 60% category B and 40% category A. For the solely behaviour experiment, category A was always comprised of the orange and green balls, and B always comprised of the pink and blue balls. In the EEG experiment, the two colours used to form each category varied between participants, but was always consistent for each individual participant for the duration of the experiment. However, for ease of presentation, the task will be described as though all participants had the same category-colour pairs as in the behaviour version

(i.e. as though for all 52 participants, category A was green and orange, and category B was blue and pink).

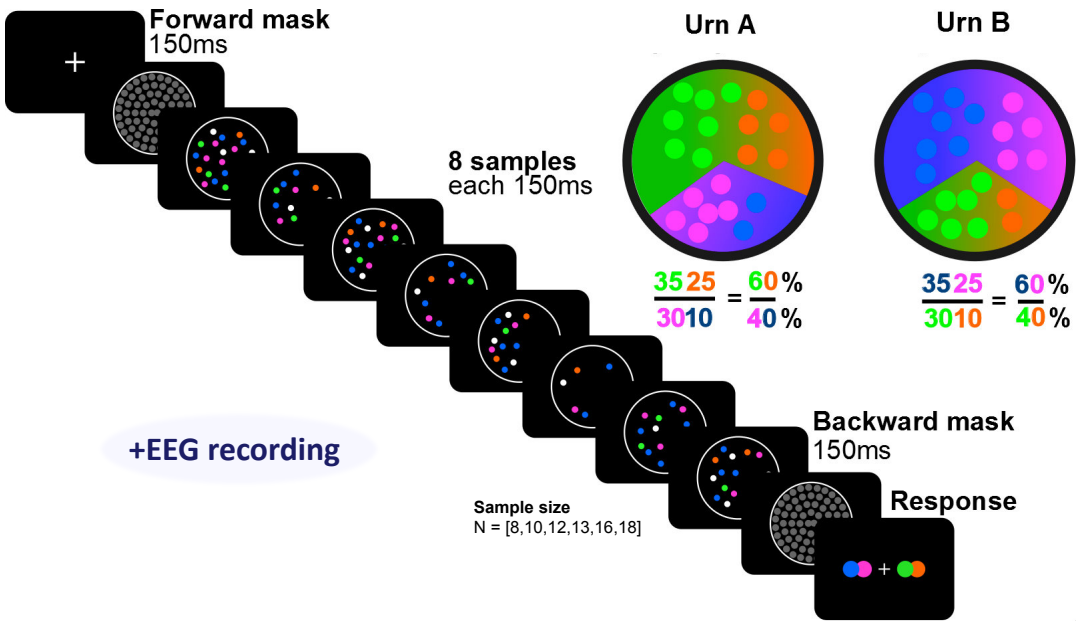

Figure 3.1. Experimental protocol. Each trial commenced with a blank screen followed by a forward mask consisting of grey dots in every possible location of the target dots. Eight draws (with replacement) were made on each trial from one of two virtual "urns": urn one had a 60:40 predominance of category A:B, and urn two had a 60:40 predominance of category B:A. The two categories (A and B) were themselves comprised of two subcategories (colours): in the example illustrated here, 'category A' consisted of green and orange balls, and category B consisted of blue and pink balls. (The colours that comprised a category differed between participants, but remained constant for each individual for the entire experiment). Crucially, the ratio of the two subcategory colours within each category was uneven, as illustrated, such that in both categories, information from one subcategory was more prevalent than information from the other. Participants' task was to indicate via key press whether the balls were drawn from the predominantly category A urn, or the predominantly category B urn (thus the asymmetric subcategory ratio is irrelevant to the optimal task solution, which treats all within-category information as equivalent). The draws from the urn were represented on screen as coloured dots within a circular aperture. Each screen also displayed between 1 and 8 white distractor dots. A backward mask identical to the forward mask and a blank screen preceded the choice prompt screen. Participants received auditory feedback via a high (800Hz) tone for correct and a low (400Hz) tone for incorrect responses. Importantly, urns were counterbalanced such that the more prevalent colour within a category changed pseudorandomly and independently for each of the two categories that comprised an urn.

Crucially, the proportion of the two colours that comprised each category was not 50:50 – within the category (A or B) that comprised 60% of the urn on a given trial, the ratio of the less prevalent colour to the more prevalent colour was 5:7, and within the category that comprised 40% of the urn this ratio was 1:3. The colour within each category that was more or less prevalent (henceforth the "dominant" and "non-dominant" subcategories respectively) on a given trial also varied pseudorandomly from trial to trial. (Thus, in total there were eight possible colour to subcategory assignments, 4 per urn – see **Table 3.S1**).

On each trial, eight draws (with replacement) of 8, 10, 12, 14, 16 or 18 balls were made from the relevant urn. The coloured balls were represented to participants on the screen as filled dots within a circular aperture (**Figure 3.1**). After viewing each of the 8 samples, the participants' task was to indicate via keyboard press whether the balls were drawn from the predominantly category A (e.g. green and orange) urn, or the predominantly category B (e.g. pink and blue) urn. Note that participants were not explicitly instructed about the uneven within-category colour distribution.

The stimuli were presented on a 17-inch CRT monitor with resolution of 1024 x 768 pixels and a refresh rate of 60Hz, using the Psychophysics 3 Toolbox (Brainard, 1997) running in MATLAB (Mathworks, Nantick, USA). Participants sat approximately 70cm from the screen in a quiet, darkened room. Before the experiment began, the task was explained to participants via an instruction screen that contained a visual cue indicating the ratio of balls in the 2 categories, with experimenter clarification if requested. Each trial began with a white central fixation cross (presented for 1000ms), a blank screen (1000ms) and then a forward mask followed by 8 sequentially presented sample screens and a backward mask, each presented for 150ms with a jittered inter-stimulus interval (ISI) of 100-200ms (average 150ms). There was then a jittered pre-response interval (in which participants saw a blank screen) of between 450 and 550ms, before a screen that prompted participants for their choice, which timed out with a 'too slow' message if response was overly torpid (>3 seconds).

Each sample screen consisted of an array of orange, green, blue, pink and white dots, each 20 pixels in diameter, that were randomly spatially distributed among 71 possible locations, with minimum separation of 10 pixels, within a circular aperture of 150 pixel radius. The forward and backward masks consisted of grey dots at each one of these possible 71 locations. They were designed to reduce primacy and recency effects (behaviourally and neurally) and participants were instructed to ignore

them. The number of coloured dots, and the proportions of each of the four colours, was determined by drawing between 8 and 18 balls from a 'predominantly category A' urn, or a 'predominantly category B' urn, as described above. Two to eight white dots were randomly added to each sample; participants were also instructed to ignore these, and their purpose was to decorrelate low-level visual information from decision information.

Participants' task was to indicate via key press ('z' key with left hand, or 'm' key with right hand) whether balls were drawn from the predominantly category A (e.g. green/orange) urn, or the predominantly category B (e.g. pink/blue) urn. In the EEG experiment, the assignment of key code to urn was counterbalanced across participants. Participants received auditory feedback on every trial: a high tone (800Hz) indicated a correct answer and a low tone (400Hz) indicated an incorrect answer.

Statistically optimal solution

Let us denote the number of orange and green balls (i.e. the subcategories forming category A) in sample *k* as $A1_k$ and $A2_k$, and the number of pink and blue balls (the subcategories that comprised category B) in sample *k* as $B1_k$ and $B2_k$ respectively. The total number of balls on each sample is denoted n_k. The probability of drawing a green or orange ball from the predominantly orange/green urn (urn A) was *p*, and the probability of drawing a green or an orange ball from the predominantly pink/blue urn (urn B) was 1-*p*. By design, *p* = 0.6.

The optimal solution to the task treats all within-category information (i.e. the two subcategories that comprise each category) equivalently, and is given by the integration of binomial probabilities. For any sample *k*, the likelihood that the balls are drawn from the predominantly green/orange urn (urn A) is given by:

$$p'_k = B(A1_k + A2_k, n_k, p)$$

Eq. 3.1

where *B* denotes the binomial probability distribution. The optimal decision rule (optimal DV) is defined by whether the sum of the log likelihood ratios of the balls in each sample coming from either

urn is greater or less than zero at the end of the trial. Thus, on each sample k, the optimal decision update (DU) for or against each response is quantified as the log likelihood ratio:

$$log\left(\frac{B(A1_k + A2_k, n_k, p)}{B(A1_k + A2_k, n_k, 1-p)}\right)$$

Eq. 2

Models of evidence integration

The integration of the difference of category evidence (i.e. $A1_k$+$A2_k$ – $B1_k$+$B2_k$) is formally equivalent to the binomial decision update, and provides a neat account of how a computationally complex process could be implemented in a neurobiologically plausible manner (see chapter 2). Thus, in the following, we will use the difference method to derive the decision variable for all models.

Unlike the optimal solution, we hypothesised that participants would selectively ingrate more prevalent category/subcategory information. In other words, we believed that on a given sample, $A1_k$ and $A2_k$ would not simply be summed and compared to $B1_k$ and $B2_k$, but instead information *within* a category may be selectively weighted based on its prevalence (eg $A1_k$ may be integrated preferentially over $A2_k$ if the former is more prevalent), and additionally information *between* categories may be selectively weighted. In other words, A_k may be integrated preferentially over B_k (where $A_k = A1_k + A2_k$ and $B_k = B1_k + B2_k$) if the former is more prevalent, over and above any within-category weighting. In what follows, the category that comprised 60% of the urn is referred to as the "majority" category, and the category that comprised 40% of the urn as the "minority" category, irrespective of which colours this happened to be on a given sample. Furthermore, we define the more prevalent colour within each category as the 'dominant' subcategory, and the less prevalent colour as the 'non-dominant' subcategory. Thus using this 'prevalence framing' terminology (as opposed to 'colour framing'), each sample featured information from two categories – majority and minority – and a total of four subcategories: majority dominant, majority non-dominant, minority dominant and minority non-dominant.

We defined five key models with various degrees of selectivity: an near-optimal model that was affected only by noise (i.e. one with no selectivity, referred to as the 'logistic model'), a leaking model, which incorporated an information decay (no selectivity), a category-selective model which preferentially integrates majority over minority information, a subcategory-selective model which preferentially integrates dominant over non dominant information, and a doubly-selective model which preferentially integrates majority over minority information, dominant information over non dominant information, and features an interaction between majority and dominant information.

In their implementation, these models were nested, differing according to which of a set of five parameters were free or fixed. The first parameter, *sigma* σ, is the inverse slope of the logistic function, and is proportional to the level of "late" noise in the decision, with higher values denoting greater noise. The second parameter reflected a bias, ϑ, and was the inflection point of this slope. ϑ allowed us to capture trivial choice patterns that may have been caused by a response hand bias. *Gamma* γ reflected an information leak. Lower values of gamma denote a greater rate of (exponential) information decay over the course of a trial. ω_{cat} down-weighted all information that was not in the majority category (the equivalent of prioritising information in the majority category). Similarly, ω_{subcat} down-weighted all information that was not in the dominant subcategories, irrespective of whether or not it was in the majority or minority category. The model with both a ω_{cat} and a ω_{subcat} parameter (the doubly-selective model) allowed an interaction in the weighting of category and subcategory membership. See table 1 for details of which models contained which free parameters. If a parameter was not free in a given model it was fixed to 1, i.e. it had no effect.

Table 3.1: free parameters per model

Model	Sigma σ	Bias ϑ	Gamma γ	Selective weight ω_{cat}	Selective weight ω_{subcat}
Logistic model	✓	✓			
Leaking model	✓	✓	✓		
Category-selective	✓	✓	✓	✓	
Subcategory-selective	✓	✓	✓		✓
Doubly-selective	✓	✓	✓	✓	✓

We defined two further models that are detailed in the supplementary information: one which simply integrated the information in proportion to its magnitude, but without respect to its category membership, and another that integrates selectively only the most prevalent subcategory on each sample. These were post-hoc additions to ensure our results could not be explained trivially by alternative mechanisms.

Defining information prevalence

Prevalence – i.e. which colours comprised the majority category, and which colours comprised the dominant subcategories – was defined according to the 'local winner', i.e. according to the locally prevalent information on a given sample:

$$A1_k + A2_k > B1_k + B2_k \quad \rightarrow \; Majcat_k = A_k, \quad Mincat_k = B_k$$
$$A1_k + A2_k < B1_k + B2_k \quad \rightarrow \; Majcat_k = B_k, \quad Mincat_k = A_k$$

For cases in which A and B were equally prevalent, the two categories were given identical weighting of $(1 + \omega_{cat})$ /2.

The two dominant subcategories on each sample were defined in an analogous way:

$$A1_k > A2_k \quad \rightarrow \; Dom_k = A1_k, \quad Nondom_k = A2_k$$
$$B1_k > B2_k \quad \rightarrow \; Dom_k = B1_k, \quad Nondom_k = B2_k$$

The opposite assignment was given for the opposite inequality, and subcategories within a given category were given identical weighting if the two were equally prevalent of (1+ω_{subcat}) /2

This allows whichever colour constitutes each category or subcategory to vary between samples (*k*) within a trial (e.g. the majority dominant colour may be A1 on sample *k*, but B1 on sample *k*+1, within the same trial *t*). Note that this does not reflect the generative structure of the task, which assigned colours to subcategories consistently across all 8 samples in a trial and was not directly observable by participants. The definition is based purely on the momentary information available on a sample-to-sample basis, without reference to the information from previous samples, and as such implies that any selective integration occurs based on the transient rather than cumulative information available.

Decision update

On each sample k, the decision update (DU) for the models was calculated by computing the difference between category A information and category B information, after any relevant weighting. On each sample, minority category information, and non-dominant subcategory information, were downweighted by ω_{cat} and ω_{subcat} respectively, whilst the information in the majority category and dominant subcategories was fully integrated. In the illustrative equation that follows, A is the majority category, and B the minority category (if the assignments were the opposite, the ω_{cat} would apply to A instead). Furthermore, in this illustrative equation, A1 and B2 are the dominant subcategories in their respective categories (and if the assignments (independently for each category) were the opposite, the ω_{subcat} would apply to the other subcategory). Thus, on each sample:

$$\alpha A_k = A1_k + \omega_{subcat,k} A2_k$$

$$\alpha B_k = \omega_{cat,k}(B1_k + \omega_{subcat,k} B2_k)$$

Where α denotes that relevant weighting has been applied. The decision update was then calculated as follows:

$$DU_{model,k} = \alpha A_k - \alpha B_k$$

The decision variable (DV) represented the accumulative evidence on each trial, with a leak of information γ incorporated:

$$DV_{model} = \sum_{k=1}^{8} DU_k \cdot \gamma^{8-k}$$

Late noise was simulated by passing the resulting DV through a logistic choice function for each subject, with inverse slope σ and the inflection point ϑ as free parameters in all models (see above), in order to generate choice probabilities.

Finally, simulated model choices were generated by comparing these choice probabilities (CP) to uniform random numbers between 0 and 1, 'category A' was selected if CP>random number.

Parameter fitting and model comparison

To find the free parameter values for each model that allowed it to account best for human performance, we used gradient descent (Matlab function fmincon) on the five models to find the parameters that meant the model fit most closely to the human data. We found the negative log likelihood (-LL) – the logarithm of the model's probability of choosing the same option as the humans – via gradient descent on even-numbered trials only. The best fitting parameter values (i.e. those that generated the minimum –LL, denoting greater similarity between human and model choices) were then selected and applied to the odd-numbered trials. The -LL derived in turn from the application to the left-out trials was used as the critical value for each model. For optimisation , the sigma parameter was given a lower limit of 1 and no upper limit. The gamma parameter had a lower limit of 0 and an upper limit of 1. The weighting parameters all had a lower limit of 0 and an upper limit of two, to account for the situation that a participant in fact over-weighted the relevant sub/category. The bias parameter could range from -10 to 10. The resulting crossfit -LLs were the quantity entered into the VBA toolbox for formal Bayesian model selection.

Our second method of comparing model and human data was qualitative comparison between human and model performance (see below for further details). For the qualitative analyses, we simulated model choices using best-fitting parameters generated from the gradient descent procedure applied to the whole dataset (i.e. with no leave-one-out procedure), in order to generate more robust parameter estimates.

Behavioural analyses

We compared human and model performance qualitatively in a number of different ways. Firstly, we plotted psychometric curves to illustrate the probability of responding 'category A' as a function of the optimal DV (for humans), and for the model-predicted DVs and corresponding simulated model data for each model.

Next, to address our key questions, we used probit regression to assess how the influence of information within each of the subcategories (A1, A2, B1, B2) was influenced by that information's category or subcategory membership. For example, if there is selective integration of the majority category, the influence of the number of balls in subcategories A1 and A2 on choosing category A is predicted by the selective integration model to be enhanced if category A is the majority (i.e. "locally winning") category on a given sample. Thus we took the number of balls in each of the four colours – A1, A2, B1 and B2 – as regressors predicting p(choose category A), as well as the interaction of the number of balls A1, A2, B1 and B2 with a quantity that was +1 if the relevant colour formed the majority category on that trial, and -1 otherwise. We summed these predictors across the eight samples in a trial, and used them to predict P(choose category A):

$$p(choose\ A) = \phi\left[b + \sum_{k=1}^{8}\beta 1_k \cdot A1_k + \sum_{k=1}^{8}\beta 2_k \cdot A2_k + \sum_{k=1}^{8}\beta 3_k \cdot B1_k + \sum_{k=1}^{8}\beta 4_k \cdot B2_k + \sum_{k=1}^{8}\beta 5_k \cdot A1_k \cdot I_k + \sum_{k=1}^{8}\beta 6_k \cdot A2_k \cdot I_k + \sum_{k=1}^{8}\beta 7_k \cdot B1_k \cdot I_k + \sum_{k=1}^{8}\beta 8_k \cdot B2_k \cdot I_k\right]$$

where ϕ denotes the standard normal cumulative distribution function, and I denotes the indicator quantity that is 1 if the colour is in the majority category on a given sample and -1 If it is in the minority category.

Similarly, if there is selective integration of the dominant subcategories, the influence of the number of balls in, for example, subcategory A1 will be enhanced if $A1_k$ is the dominant subcategory (relative to $A2_k$) on a given sample, and the same holds for the dominant subcategory in that sample from category B. To test this, we performed an identical regression, but the indicator quantity I denoted whether or not $A1_k$ and $B1_k$ were the dominant subcategories on a given sample with +1 or -1 respectively.

EEG acquisition and preprocessing

A neuroscan EEG system with NuAmps digital amplifiers was used to record EEG signals from 64 Ag-AgCL scalp electrodes, with a further four electrodes used in a bipolar montage as horizontal and vertical EOGs, and a further two electrodes located on the mastoids as a reference. The electrode impedances of all functioning electrodes were kept below 10 kΩ The EEG signals were recorded at a 1kHz sampling rate, and were high pass filtered online at 0.1Hz.

The data were preprocessed using the EEGLAB toolbox for MATLAB (Delorme & Makeig, 2004) alongside custom scripts. The data were down-sampled to 250Hz and then bandpass filtered between 1 and 40Hz. Electrodes that consistently malfunctioned or produced highly noisy signal (based on visual inspection of the data) were interpolated, the data were re-referenced to the average signal and then epoched from 1000ms before the first sample onset to 5000ms afterwards (thereby covering the entire trial). Epcohs were visually inspected to remove trials containing non-stereotypical artifacts. An independent component analysis was then conducted using EEGLAB, and the resulting components were visually inspected to remove artifacts (particularly blinks). Finally, the resulting data were re-epoched from 500ms before to 700ms after each sample.

EEG analyses: Encoding

The use of nested models meant that behaviour predictions at the trialwise (choice) level were often similar, despite implying quite different underlying mechanisms. However, we were able to gain a more in-depth understanding of the mechanisms at play at each timepoint by investigating the concurrent EEG activity. Particularly, we were interested in which models best explained the variance in the neural signal. Thus, to this end, rather than calculating ERPs, as in Chapter 2 we used a general linear modelling approach (at the single sample level) that allowed us to ascertain how the variability in the EEG signal was influenced by various sources of information on each individual sample, at each timepoint over a period spanning -100ms-700ms from stimulus onset.

We began by performing an identical regression to that described for behaviour, but this time predicting the average EEG signal in the occipitoparietal (PO7, PO5, PO3, POz, PO4, PO6, and PO8) region rather than predicting choice. We focused on this region based on its previous associations with evidence integration (e.g. (Roitman & Shadlen, 2002)). The eight predictors were entered competitively in the same regression in order to ensure that they competed for unshared variance, and the decision information (the number of dots in A1, A2, B1 and B2) from the previous and following samples was included in the regression as "nuisance" covariates. This step enabled us to ensure that the parameter estimates were not influenced by information from the adjacent samples that overlapped in time with the epoch. Thus, the resulting parameter estimates indicated the unique correlation at each timepoint following sample onset between the neural signal in the given region, separately for the eight predictors.

We expected all decision-relevant information to have some impact neurally, but our key question of interest was whether or not prevalent category and/or subcategory information was encoded with significantly greater gain (i.e. steeper slope of encoding) than less prevalent information, and if so which model(s) predicted this pattern. To this end, parameter estimates from the regressions were then averaged across samples and entered into second level statistical analyses in order to ascertain whether the encoding of the decision information (the number of dots in each subcategory) was enhanced according to whether the information was in the majority category and/or in the dominant

subcategory. Regions of time and space (i.e. electrodes) at which these betas deviate significantly from zero across all participants indicated that these effects are reliably encoded in neural signals. In order to correct for multiple comparisons across time, we used a nonparametric cluster correction technique with an alpha of 0.05 (Maris & Oostenveld, 2007).

Control analyses

We wanted to establish that any asymmetric weighting due to category membership was not a by-product of the task design or response contingencies. The existence of any systematic biases at the processing level, rather than the response level, could be problematic for overall interpretation of our results. To this end, we first assessed the extent to which participants were biased to respond with their dominant hand (the structure of the task meant that the participant's right hand would signal the correct answer exactly 50% of the time). Next, we assessed performance (i.e. percent correct) separately for trials on which the correct answer was 'A' or 'B'. Following this, we conducted a signal detection style analysis by calculating d' and the criterion separately for each category, to ensure that any differences seen in performance between the two categories were a result of a criterion (response) bias, rather than a sensitivity (and therefore potentially mechanistic) difference. These results motivated the inclusion of a bias term in the models.

3.3 Results

Behaviour analyses

Results are collapsed across the two versions of the experiment (solely behaviour, and EEG) in all behaviour analyses reported here. Throughout, model performance is generated using the parameterisation that minimised the -LL in the gradient descent procedure, as described above. Participants chose the correct urn on 78.58 ± 6% of trials, with a response timeout on 0.001% of trials and an average reaction time of 0.65s on all other trials. All models performed very similarly to humans (**table 3.2**).

Table 3.2: overall performance (% correct) per model

Model	Average score from simulated choices
Logistic model	79.32 ± 6%
Leaky model	79.01 ± 6%
Category-selective model	79.41 ± 6%
Subcategory-selective model	77.94 ± 6%
Doubly-selective model	77.44 ± 7%

We began by plotting human psychometric functions (choice probabilities as a function of binned DV). The relationship between the DV and choice was as expected: the choice probability grew with the magnitude of the DV (**Figure 3.2b**). A small response bias was evident in the data that may reflect greater tendency for participants to respond with their dominant hand.

Next, we conducted a probit regression, taking the number of dots across a trial from each of the four subcategories – A1, A2, B1 and B2 - as well the interaction between these quantities and an indicator variable that was 1 if the relevant subcategory was one of the two majority subcategories, and -1 otherwise, to estimate whether the impact that information had on p(choose category A) was

influenced by whether or not it was in the majority category. The results showed that whilst, as expected, the information in the two subcategories that comprised category A positively predicted a category A choice, and the information from the two subcategories in category B negatively predicted a category A choice, there was no interaction of this impact with majority category membership. In other words, counter to the predictions of the standard selective integration model, locally winning category information was not disproportionately favoured during decisions. (**Fig 3.2c**).

When we repeated this regression to assess whether the dominant subcategory membership influenced the manner in which information was processed, we saw an intriguing pattern of results: in contrast to the majority information, if information (A1, A2, B1, B2) was in a dominant subcategory on a given sample, it had a significantly greater impact on choice, as shown by the interaction terms in figure **3.2d.** This novel pattern of results is indicative of selective integration of information in the locally winning subcategory, but not the locally winning category.

Models of evidence integration: qualitative comparisons

A model that can aptly describe human data must capture qualitative patterns of performance (Palminteri, Wyart, & Koechlin, 2017) as well as quantitative effects. We took the parameter values from the gradient descent procedure that minimised the -LL values for five models and used these to generate model choices (see method). We repeated the behavioural analyses above, but this time predicted model (rather than human) choices, in order to compare the two qualitatively. All models mirrored the human psychometric curves closely (**figure 3.2b,** purple dots, **supplementary figure 3.1S**, coloured dots).

The model performance from the two regression analyses was more diagnostic of its ability to account for human data: whilst all models mimicked the behaviour effect that showed that majority category membership lead to no greater impact on choice than minority category membership, only the subcategory-selective and the doubly-selective models were able to capture the phenomenon whereby dominant subcategory information had an enhanced impact on choice relative to non-dominant subcategory information (**figure 3.2 c and d,** purple dots, **supplementary figure,** coloured dots). This qualitative effect was reflected quantitatively by the fact that the four beta weights for the interactions differed significantly from zero for humans (all $|t|>3.13$, all $p<0.001$) as well as for the

subcategory-selective model (all $|t|>5.84$, all $p<0.001$), where the critical alpha level corrected for multiple comparisons is 0.0125.

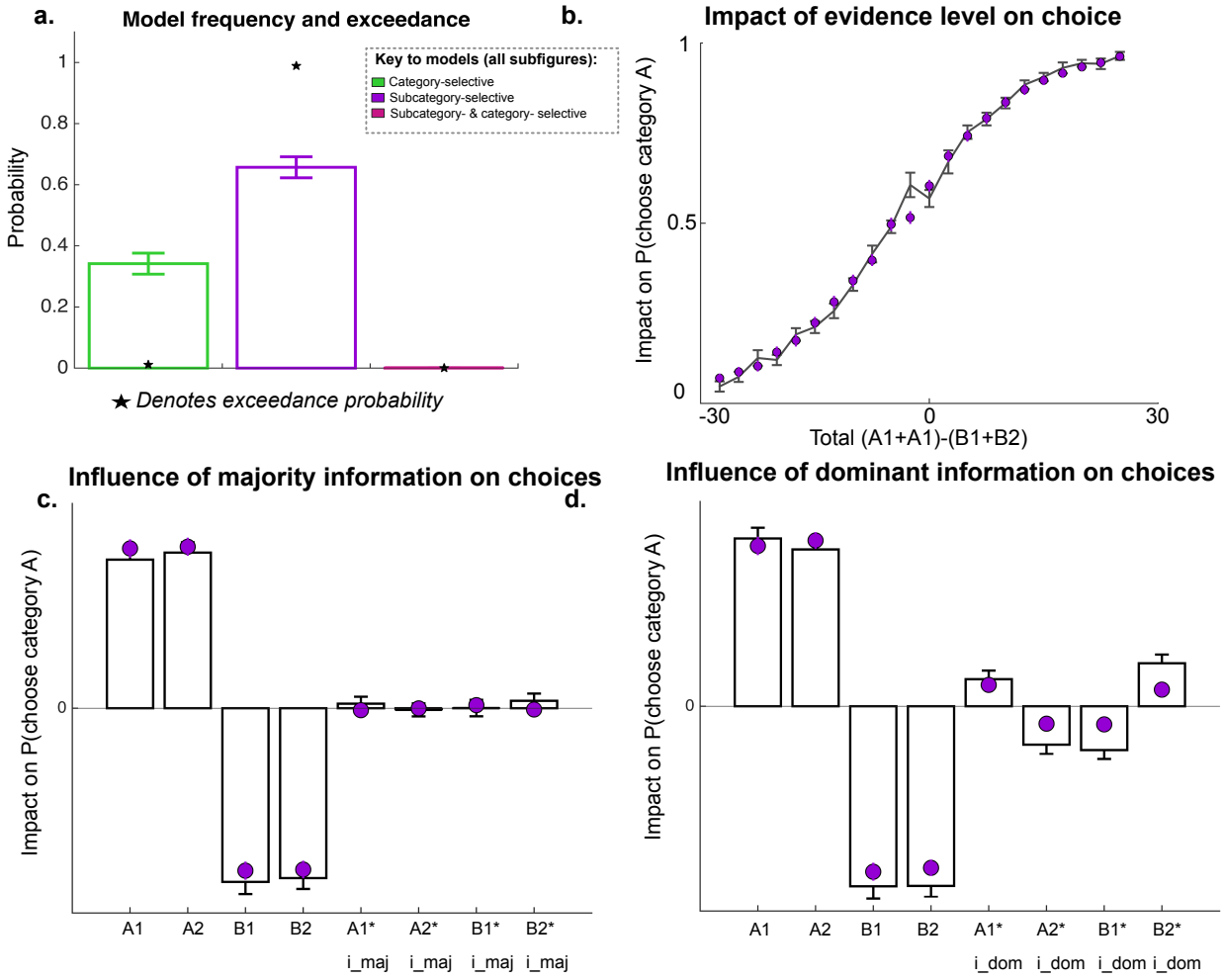

Figure 3.2. Model fitting and behaviour results. (A). Quantitative model comparison: expected frequency (bars) and exeedance probability (stars) derived from Bayesian model comparison for the category-selective, subcategory-selective and doubly-selective models. For both measures, higher values suggest that the model provides a better explanation of the human data to which it is being fit. Our focus on these three models was due to the fact that only models with a weighting parameter ω were able to account for qualitative results, as shown in the other panels of this figure and supplementary figure 3.1S. (B). Probability of humans (lines) and the subcategory-selective model (purple dots) choosing category "A", as a function of the value of the optimal decision variable (i.e. no selectivity, bias or noise). (C). Impact (beta coefficient) of humans (bars) and the subcategory-selective model (purple dots) choosing 'category A' of information in the four subcategories (A1, A2, B1 and B2; first four bars/dots), and the interaction between this quantity and whether or not the information formed the majority category (final four bars/dots). i_maj was 1 if the relevant subcategory was in the

majority category, and -1 otherwise. The lack of interaction shows that the impact of the information in A1, A2, B1 and B2 was not affected by whether or not it was in the majority category, counter to what the selective integration model would predict. (D). Impact (beta coefficient) of humans (bars) and the subcategory-selective model (purple dots) choosing 'category A' of information in the four subcategories (A1, A2, B1 and B2; first four bars/dots), and the interaction between this quantity and whether or not the information was in one of the two dominant subcategories (final four bars/dots). Unlike the majority regression results, the results here show that information has an enhanced impact on choice if it in a dominant subcategory. i_dom was 1 if A1 or B1 were the dominant subcategories, and -1 otherwise, thus (for example) A1*i_dom reflects the increase of impact on choice that A1 has when it is the dominant subcategory, and A2*i_dom relects the decrease of impact A2 has on choice when A1 is the dominant subcategory.

Quantitative comparison

Having thus established that only models that contained at least one weighting parameter were able to capture key qualities of the human data, we compared the three models that featured this weighting parameter – the category-selective, subcategory-selective and doubly-selective models – using the VBA toolbox. The VBA toolbox calculates the Bayesian posterior probability of each model, using a random effects procedure that allows the best fitting model to differ between participants. The quantities of interest here are the exceedance probability – the measure of how likely it is that each model is more frequent than all other models in the comparison set – and the model frequency, which denotes the proportion of participants in the cohort for whom a given model had the highest posterior probability of explaining their data. Because we used a cross-fitting procedure, we were able to enter the minimum -LL from the best fitting parameter values derived from gradient descent without the need to apply further penalty. The results of the model comparison strongly favoured the subcategory-selective model over the category- and doubly-selective models (**figure 3.2a**): the exceedance probabilities and frequencies for the category-, subcategory- and doubly-selective models were 0.01, 0.99 and 0 respectively, and the expected frequencies (with chance at 33%) were 34%, 66% and 0% respectively, indicating a strong quantitative advantage for the subcategory-selective model.

As described, models with a weighting parameter ω_{cat} and/or ω_{subcat} implemented the selectivity by *down*-weighting all information that was *not* in the majority or dominant subcategories by the value that the parameter took, and fully integrating the majority/dominant information. Models that assume individuals selectively integrate information (either at the category level, the

subcategory level, or both) would predict that the values of the weighting parameters that lead to the best fit to human data would be lower than 1, as higher than 1 implies anti-selectivity, i.e. that the majority/dominant information would itself be processed with *lower* gain relative to the other information. Nonetheless, we allowed the values in the gradient descent search to range between 0 and 2, to allow for this possibility. We conducted a t test against 1 for the best fitting weight parameter values (ω_{cat} and/or ω_{subcat}) derived from the gradient descent search (see table 3.3 for the average parameter values). The average value of ω_{cat} in the category-selective and doubly-selective model was 0.94 in both models, and this was, as expected, significantly lower than 1 across the cohort: *t*(51)=-2.301, *p*=0.04 and t(51)=-2.177, p=0.034 respectively. The same pattern was seen for subcategory selectivity: the average value of ω_{subcat} in the subcategory-selective model and the doubly-selective model was 0.77 and 0.75, t(51)=-4.121, p<0.001 and t(51)=-4.07, p<0.001 respectively. The best fitting parameter values for all parameters for all models, averaged across participants, were as follows:

Table 3.3: best fitting parameter values

Model	Sigma σ	Bias ϑ	Gamma γ	Selective weight ω_{cat}	Selective weight ω_{subcat}
Logistic model	**10.61**	**-3.86**	**-**	**-**	**-**
Leaking model	**6.74**	**-2.71**	**0.87**	**-**	**-**
Category-selective	**7.05**	**-2.88**	**0.86**	**0.94**	**-**
Subcategory-selective	**6.05**	**-2.45**	**0.87**	**-**	**0.77**
Doubly-selective	**6.26**	**-2.54**	**0.87**	**0.94**	**0.75**

Late noise simulations

Selective integration at the category level has previously been shown to protect against the negative influence of late noise on decision accuracy. Here, in a multi-feature paradigm, we found evidence for

a new type of selective integration, as although the parameter values were lower than 1 for ω_{cat}, the best fitting model integrated selectively only at the subcategory- rather than category-level. Therefore we wanted to ask whether there is an advantage to selective integration at the subcategory level, as occurs in our best fitting model, in the face of late noise?

To answer this, we generated choice probabilities (i.e. p(correct)) for the logistic model, which has no weighting, the category-selective model, and the subcategory selective model, at different levels of selectivity and different levels of late noise (with all other possible free parameters fixed to have no impact).

As a sanity check, we first showed that in the logistic model, the only impact on correct choices was noise: as simulated late noise increased, simulated performance decreased (**figure 3.3,** upper panel). In line with previous work, the category-selective model was more robust than the optimal model in the face of late noise: at the higher levels of noise, performance declines as selectivity decreases (lower weight values denote higher selectivity), **figure 3.3,** middle left panel. At lower levels of noise, category-selective model performance is relatively stable across levels of selectivity, though weighting values greater than 1 (i.e. relative *down*-weighting of the majority category) harm performance at every value of noise, as we would expect. However, the simulated values for the subcategory-selective model painted the opposite picture: at all levels of noise, higher selectivity (lower weight values) appeared to be detrimental to performance (**figure 3.3,** middle right panel).

However, applying down-weighting will have a general impact on performance as it lowers the overall value of the DV, therefore decreasing the signal to noise ratio. Thus the appropriate comparison to see the effects of *selectivity,* rather than general harm that may be caused by decreasing the size of the DV, is to compare performance of a selective model with an "equivalent-gain" model: one that down-weights all information equivalently (non-selectively) by a factor that brings the total decision information onto the same scale as the selective integration models. Comparing the performance of an equivalent gain model with that of the category-selective model shows more clearly the advantage of the latter, and also demonstrates clearly that under any level of late noise (low or high), weighting parameter values of higher than 1 are detrimental (**Figure 3.3,** lower left panel). Interestingly, subcategory selective integration has no advantage at any level of weighting or any level of late noise

above the equivalent gain model. See the discussion section for the implications this may have on the selective integration framework.

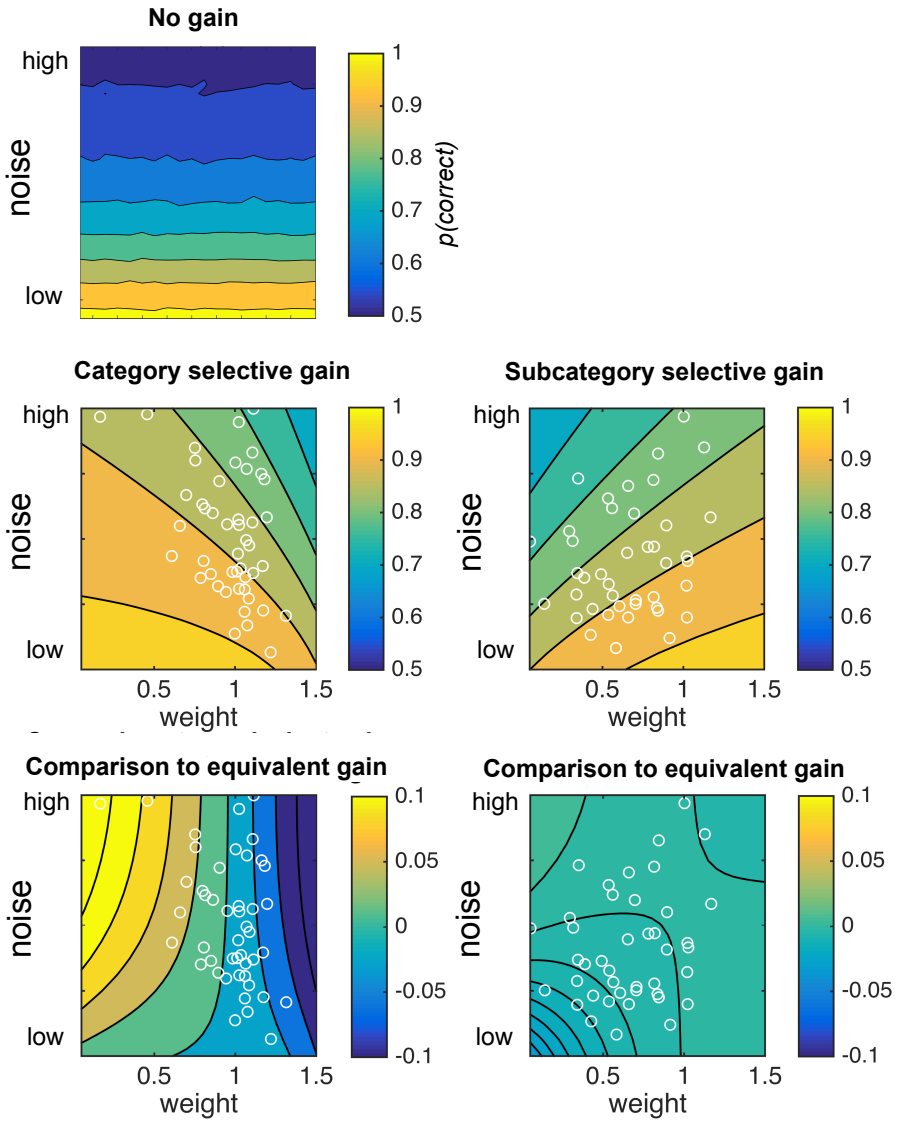

Figure 3.3. Late noise simulations. The first three subplots each predict the probability of a correct choice under varying levels of noise (y axis) and selectivity (i.e. weight, x axis). A lower weight means higher selectivity (i.e. stronger down-weighting of the less prevalent subcategory). On all subplots, white circles show the actual best fitting parameters for each individual subject. The top subplot depicts the probability of a model with no selective weighting, thus the only thing affecting performance is the level of noise, with performance decreasing as noise increases as we would expect. The left and right middle panels show the category- and subcategory-selective models respectively. With category selectivity, performance under high late noise is enhanced by higher selectivity (lower weight), but under lower noise, performance is relatively unaffected by selectivity. The opposite

pattern is seen for the subcategory selective model. However, when these selective models are compared against equivalent gain models (i.e. models that down-weight information across all subcategories such that gain is equivalent to the selective models), the relative advantage of the category-selective model is more evident: the bottom left panel shows p(correct) for the category-selective model minus p(correct) for the non-selective model with equivalent. The same is shown for the subcategory-selective model on the lower right hand panel: interestingly there is neither an advantage nor a disadvantage for subcategory selectivity relative to no selectivity at equivalent gain.

EEG analyses

The behaviour data show qualitative and quantitative support for a model that selectively integrates subcategory information on the basis of its prevalence. Rich information from neural recording at each timepoint over the course of the trial allows us to garner more in-depth support for the mechanisms that the subcategory-selective model proposes subserve the decision process.

Encoding analyses

The subcategory-selective model predicts enhanced gain for dominant over non-dominant subcategory information, but no difference in encoding for majority over minority category information. Is this reflected in the neural data? We began, as a sanity check, by regressing at each timepoint from -100ms to 700ms following stimulus onset the total information on each sample from category A ($A1_k$+$A2_k$) and category B ($B1_k$ and $B2_k$) respectively against the average EEG activity in occipitoparietal electrodes. We expected to see significant encoding of this decision information, which was indeed evident (**figure 3.4a**).

We next repeated the regression taking as predictors the information on each sample from each subcategory separately (i.e. $A1_k$, $A2_k$, $B1_k$, and $B2_k$) and the interaction of this information with an indicator that was 1 if the subcategory was in the majority category and -1 otherwise (in an identical way to the behavioural regressions, with the only difference that the information was regressed against the sample-wise signal in these EEG regressions and therefore was not summed across the trial). The encoding of the decision information in parietal regions was not affected by whether or not it was in the majority category (**figure 3.4b**). However, when the regression was repeated but this time with an indicator variable that reflected whether or not the information was from the dominant subcategory, the results showed enhanced encoding of information from the dominant subcategory

(**figure 3.4c).** This finding clearly mirrors that seen in behaviour, providing a parsimonious link between the neural and behavioural data.

That we are able to gather neural support for the effects predicted by the best fitting model is important because it verifies that, as well as behaviour, the continually evolving neural activity is closely linked to the model predictions at all stages of the unfolding trial. Furthermore, the revelation that some information is processed with enhanced gain from the outset provides support for selective integration **at the time of encoding:** if information were fully integrated but only later prioritised or demoted to have a higher or lower impact on choice, we would not expect to see this enhanced processing of dominant information at the time of encoding. To our knowledge, this is the first direct neural evidence supporting the general notion of selective integration of decision information.

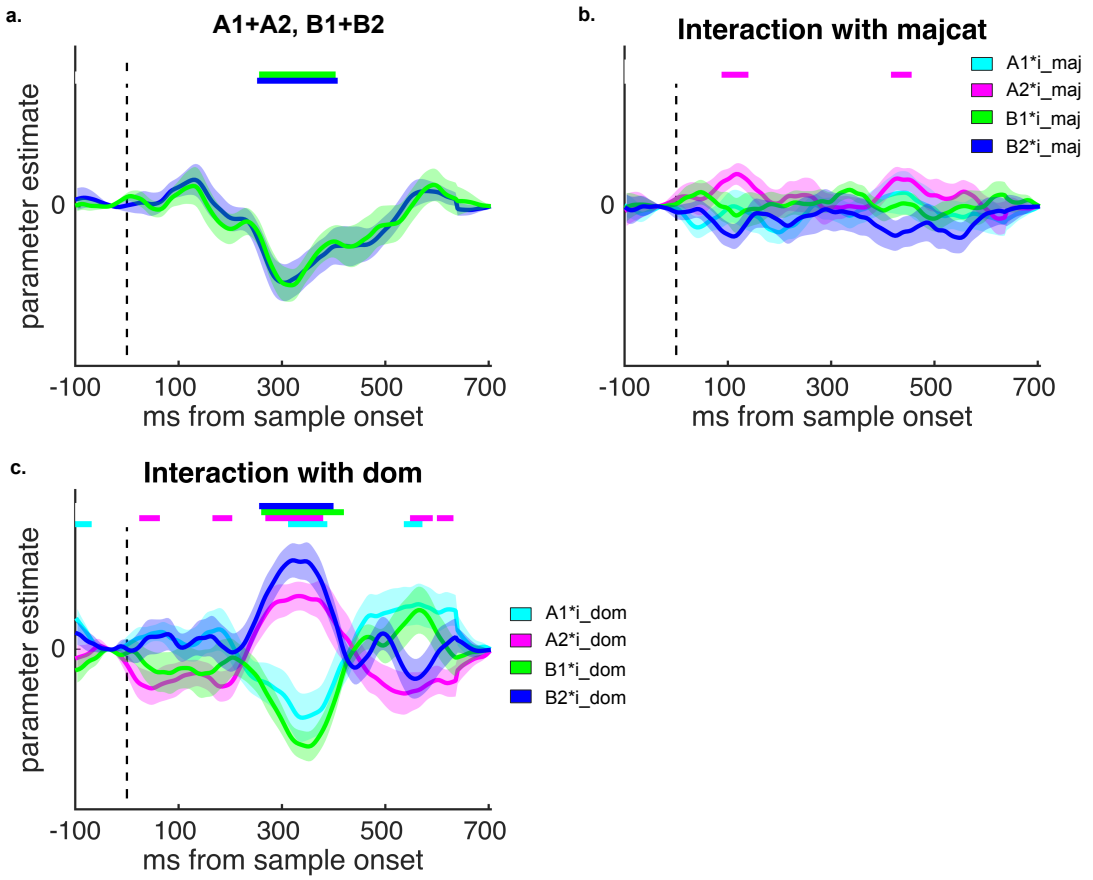

Figure 3.4. Neural encoding curves from parietal regions. (A). Relationship between EEG activity in parietal regions and the level of category A information (blue curve) and category B information (green curve). The lines at the top show periods of time at which these 'encoding curves' deviated

significantly from 0, indicating that the relevant quantity is encoded with above-chance strength. (B). The influence that being in the majority category has on the encoding of decision information in each subcategory (A1 and A2, B1 and B2). i_maj was 1 if the relevant subcategory was in the majority category, and -1 otherwise. The impact is minimal, suggesting that the encoding of decision information is minimally modulated by category prevalence. (C). The influence that being in the dominant subcategory has on the encoding of the relevant subcategory. i_dom is 1 if A1 or B1 are in the dominant subcategory, and -1 otherwise. The results show that the encoding is indeed significantly affected by subcategory membership: if the given information is in the dominant subcategory, it is processed with enhanced gain relative to the non-dominant subcategory. This pattern is predicted by the subcategory-selective model.

Control analyses

The EEG participants each saw different colour-category combinations – for example, one participant may have had category A comprising green and pink balls, with B comprising orange and blue balls, whilst the next participant may have had a category A comprising pink and orange balls, and B comprising green and blue, etc. Thus, the colour and response hand assignment was counterbalanced to the extent we could be confident there would be no resulting biases. However, the behaviour-only participants all had category A comprised of blue and pink balls, and category B comprised of green and orange balls. Furthermore, these participants always indicated a category A choice with the right hand. Therefore, as a control analysis, we investigated the nature and extent of any biases that may have arisen from this underlying structure for this group of participants.

We first saw that there was a strong right-hand response bias: the ratio of right to left hand responses (and therefore proportion of times participants chose ‘A’ over ‘B’) was 56:44. Whilst ideally the link between a response hand and a specific choice would not have existed, such a bias is not uncommon in a majority right-handed population, and is what motivated the choice to counterbalance response contingencies in the EEG version that followed. It was also the case that when the correct answer was ‘A’, participants were correct 84.6% of the time, as opposed to 71.7% of the time when ‘B’ was the correct answer. This could have been due to the response hand bias, or due to the salience of the pink and blue dots above the orange and green dots, or a combination. This bias could be problematic in the interpretation of results if it influenced the manner in which information was actually processed,

rather than just influencing behaviour at the response level. To test for this, we calculated d prime (sensitivity) and bias separately for responses for each of the categories. Fortunately, the results showed identical sensitivity (d'=1.59 for both 'A' and 'B'), but very different criteria: for 'A' (right hand) responses participants adopted a liberal threshold of 0.7, whereas for 'B' (left hand responses) people were much more conservative with a criterion of 1.43. Thus overall, although there are response-hand and potential colour biases present, we do not believe this influenced the nature of processing at the mechanistic level. However, to be extra cautious, we included (as described) a bias term in all models as a free parameter, which allowed all models to account naturally for this effect, and meaning that the value of other parameters should not simply be driven by spurious inconsistencies such as this response hand bias.

Summary

The model that best accounted for human performance in this multi-attribute categorisation task was one that selectively integrated subcategory information on the basis of its prevalence. The predictions of this model captured qualitative patterns seen in human data, such as the enhanced impact the decision information had on choice if it formed part of the dominant subcategory. Neurally, key signatures of this model were evident: dominant subcategory information was processed with enhanced gain. Counter to our predictions, we saw no selective integration at the category level, meaning that our findings, whilst clearly linked to previous work on this topic, showed a qualitatively different phenomenon. Unlike selective integration at the category level, we found no evidence that suggested subcategory selectivity alone had a beneficial effect in protecting against the influence of late noise.

3.4 Discussion

Decisions are susceptible to corruption from external noise. A canonical viewpoint holds that human perceptual decisions take into account all available evidence in a procedurally optimal manner, and choices are limited only by this uncontrollable source of variance (Ma, Navalpakkam, Beck, Berg, & Pouget, 2011; Navalpakkam, Koch, Rangel, & Perona, 2010). However, recent work suggests that factors beyond susceptibility to such 'peripheral' noise are required to explain variability in human choices (Drugowitsch et al., 2016) and furthermore that "early noise", occurring at the stage of sensory processing, and "late noise", occurring at the stage of decision formation, have dissociable influences on the decision process (Herce Castañon et al., 2018; Li et al., 2017; Tsetsos et al., 2012; Tsetsos et al., 2016). Cognitive tasks, which are particularly susceptible to the effects of late noise, benefit from selective integration of information about alternatives with a momentarily higher value, as despite being counter to an optimal (full integration) strategy, it provides a robust defence against the corrupting influence of late noise on a decision whilst employing a computationally efficient strategy.

Here, we employed a multi-feature binary choice task that required the integration of multiple samples of evidence over time, in order to investigate the effects of selective integration. Whilst the choice that agents were required to make was a simple categorisation task ("Are the balls drawn from urn A or B?"), the categories themselves were comprised of more than one feature (colour). Using this design, we found a that participants selectively integrated information from the more prevalent feature within each category, but did not selectively integrate information at the category level. In other words, of the four subcategories (A1, A2, B1 and B2), A1 was selectively integrated if it was more prevalent than A2, or vice versa (and the analogous process for B1 and B2), but if A1+A2 was more prevalent than B1+B2, this did not affect the way that the total category information was processed.

Selective integration, whilst in some senses suboptimal, has been shown to confer two important advantages: firstly, it leads to more accurate decisions when late noise is higher, and secondly, it is a computationally efficient mechanism as it allows good decisions to be made without the need to integrate all the decision-relevant information fully. However, the former has only been shown in uni-

feature tasks. A particularly striking feature of the pattern of subcategory selectivity shown in this multi-feature task is that, unlike selective integration at the category level, it does not protect decisions in the face of late noise. Furthermore, although it remains beneficial from the perspective of computational load to down-weight information, the late noise simulation showed that there was no advantage (or disadvantage) to selective integration as opposed to a model that slightly lowers the processing of all information (rather than selectively), such that the computational load is equivalent. We were surprised by this finding and note that there are (at least) two interpretations. Firstly, it could be that these results are indicative of a general theme: whilst selective integration is beneficial in simple categorisation tasks, it is harmful in others, and as such people are able to behave 'optimally' in the face of late noise only under limited circumstances. Secondly, the fact that subcategory selectivity is not beneficial as late noise increases in this experiment could be a simple by-product of the task design. Specifically, subcategory assignment (i.e. which colours were in the majority category and the dominant subcategories) although consistent across a trial in the generative task structure could change several times between samples across a trial. The benefits of selective integration on each sample in the trial under such circumstances be cancelled out on other samples, whereas if there was higher consistency it would be more beneficial. This would mean that in general, selective integration at the subcategory level is advantageous, despite it not being beneficial here. Determining which (if either) of these explanations is more likely to be correct requires further empirical investigation.

After seeing that the results were well explained by a subcategory-selective model, we conjectured that people may in fact selectively integrate only the most prevalent subcategory (rather than both of the dominant subcategories), as it was an intuitive strategy to solve the task that would involve a low computational load. However, a model that behaved in this way fared badly (see supplementary information). Again this could be because the most prevalent subcategory could change several times over the course of the trial, making a momentary focus on the locally winning subcategory an un-robust strategy. Another version of this model that also failed to capture performance began with a wider focus of attention, increasing its selectivity for the prevalent subcategory as the trial progressed and the correct answer became more apparent, in the manner of a confirmation bias (Nickerson, 1998). However, whilst not supported here, it could be the case that a model that hones its selectivity over time may account for behaviour in tasks that require longer periods of integration. Such tasks

would require higher computational load that may be reduced by a confirmation-bias style process of gradually enhancing selectivity over the course of each trial.

A key advantage attributed to selective integration over full integration is this computational efficiency: good decisions can be made at lower neural computational cost. However, previously there has been no neural evidence for this conjecture: specifically, although behaviourally, effects consistent with selective integration of information have been clearly demonstrated, it is possible that these occur due to full processing of information that is only down-weighted later in the decision process (for example, via an asymmetric leak rate for locally winning pieces of information). In other words, a less computationally efficient process could give rise to identical behaviour effects. Here, the use of concurrent EEG recording provided important support for the idea that selective integration occurs at the time of presentation, as we demonstrated that from the initial processing stages, information from the dominant subcategory was processed with enhanced neural gain relative to that not in the dominant subcategories.

A substantial corpus of work shows that information that falls under the focus of attention is prioritised in neural processing (e.g. (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Maunsell & Treue, 2006; Reynolds & Chelazzi, 2004) amongst many others). Initially, we defined information prevalence (i.e. which subcategory forms the majority category and the dominant subcategories) in two ways: the first was the definition we use throughout, which assigns category membership based on 'local prevalence', i.e. the information that is momentarily more prevalent on a given sample. The second definition was according to 'global prevalence': the majority category and the dominant subcategories were defined according to which subcategory was more prevalent in the running tally of evidence in a trial on each sample. The latter definition implies that individuals calculate and refer to the cumulative decision variable (DV) at each timepoint, whilst the former implies a simpler attentional policy in favour of the momentarily winning information. The global prevalence definition did not provide a good account of human behaviour, hence the focus on models defined according to the momentary definition. That the best fitting models used a momentary attentional policy is in some ways surprising: neurally, the *cumulative* DV appears to be encoded (Mazurek, Roitman, Ditterich, & Shadlen, 2003; Roitman & Shadlen, 2002), and the importance of the running tally of evidence in determining which regions of decision space are processed with enhanced

gain has been demonstrated (Cheadle et al., 2014). However, these findings are by no means mutually exclusive: the focus on momentary prevalence sheds light on the nature of the local attentional policy that by necessity must precede later encoding of the cumulative information, and is in line with an extremely well-established body of work on the influence of selective attention on evidence integration (e.g. (Busemeyer & Townsend, 1993; Krajbich, Armel, & Rangel, 2010).

External noise is an irreducible feature of everyday decisions. A strong body of work holds that human decision makers employ a statistically optimal evidence integration strategy, corrupted only by this external noise. However, such work overlooks the impact of internal noise that occurs later in the decision process: under the influence of late noise, "suboptimal" strategies, such as selective integration of decision information, have been shown to improve accuracy. Here, we asked whether we would see the same effect in a more complex task and found that, although people clearly exhibited the selective integration of more prevalent information, this did not have the same beneficial effects against the influence of late noise. This suggests that whilst selective integration may generally benefit accuracy under late noise, and therefore as a general strategy is optimal, there are circumstances and tasks which benefit only from the reduced computational processing it allows rather than its influence against late noise.

Supplementary information

Urn structure

The urns were counterbalanced such that the subcategory membership of one colour could not be predicted in advance (e.g. it was not the case that "green" was always more prevalent than "orange") and was not predictive of the subcategory membership of colours in the alternative categories. Table 3.S1 illustrates the possible underlying urn structures, using as an example a participant who had category A comprising orange and green, and B comprising pink and blue:

Table 3.S1: Urn counterbalancing

	Urn	Colours	Proportions	Category	Subcategory	
1	A	A1 (e.g. Green)	35%	60%	Majority	Dominant
		A2 (e.g. Orange)	25%			Non-Dominant
		B1 (e.g. Pink)	30%	40%	Minority	Dominant
		B2 (e.g. Blue)	10%			Non-Dominant
2	A	A1 (e.g. Green)	25%	60%	Majority	Non-Dominant
		A2 (e.g. Orange)	35%			Dominant
		B1 (e.g. Pink)	30%	40%	Minority	Dominant
		B2 (e.g. Blue)	10%			Non-Dominant
3	A	A1 (e.g. Green)	35%	60%	Majority	Dominant
		A2 (e.g. Orange)	25%			Non-Dominant
		B1 (e.g. Pink)	10%	40%	Minority	Non-Dominant
		B2 (e.g. Blue)	30%			Dominant
4	A	A1 (e.g. Green)	25%	60%	Majority	Non-Dominant
		A2 (e.g. Orange)	35%			Dominant
		B1 (e.g. Pink)	10%	40%	Minority	Non-Dominant

		B2 (e.g. Blue)	30%			Dominant
5	B	A1 (e.g. Green)	30%	40%	Minority	Dominant
		A2 (e.g. Orange)	10%			Non-Dominant
		B1 (e.g. Pink)	35%	60%	Majority	Dominant
		B2 (e.g. Blue)	25%			Non-Dominant
6	B	A1 (e.g. Green)	30%	40%	Minority	Dominant
		A2 (e.g. Orange)	10%			Non-Dominant
		B1 (e.g. Pink)	25%	60%	Majority	Non-Dominant
		B2 (e.g. Blue)	35%			Dominant
7	B	A1 (e.g. Green)	10%	40%	Minority	Non-Dominant
		A2 (e.g. Orange)	30%			Dominant
		B1 (e.g. Pink)	35%	60%	Majority	Dominant
		B2 (e.g. Blue)	25%			Non-Dominant
8	B	A1 (e.g. Green)	10%	40%	Minority	Non-Dominant
		A2 (e.g. Orange)	30%			Dominant
		B1 (e.g. Pink)	25%	60%	Majority	Non-Dominant
		B2 (e.g. Blue)	35%			Dominant

Note that the actual colours paired to form the categories varied between participants; (see methods). The behaviour only version featured only urns 1, 2 5 and 6, thus was still counterbalanced beyond reasonable possibility that participants could predict category membership, but to a lesser extent than in the EEG version.

Additional models

We defined two further models to those described in the main results. The first was designed to ensure that any apparent selective integration effects could not be explained trivially by a model that was simply sensitive to the magnitude of evidence favouring each subcategory, without reference to its sub/category membership. The proportion model weighted each colour based on the proportion

of the sample it represented. As in the main text, let $A1_k$ and $A2_k$ represent the number of dots in the two respective subcategories that comprised category A on sample k, and let n_k represent the total number of dots in sample k:

$$Proportion\_A1_k = A1_k / n_k$$

$$\alpha A1_k = \left(A1_k + \left(Proportion\_A1_k \cdot A1_k \cdot \omega_{proportion}\right)\right)$$

Where α denotes that weighting has been applied, and with an analogous process for each of the other three subcategories (A2, B1 and B2). The DU and DV were then calculated as follows:

$$DU_{proportion,k} = (\alpha A1_k + \alpha A2_k) - (\alpha B1_k + \alpha B2_k)$$

$$DV_{proportion} = \sum_{k=1}^{8} DU_{proportion,k} \cdot \gamma^{8-k}$$

Where ϒ denotes an information leak. Thus the proportion model contained just two free parameters, one weighting parameter and a leak parameter.

The second additional model was identical in structure to those in the main text but featured "enhanced" selective integration, selectively integrating only the majority dominant subcategory and down-weighting everything else by a single weighting parameter ω with an information leak ϒ.

For both models, choices generated in an identical manner to the process described in the main text. Both models were rendered inferior by the VBA model comparison process described in the main text. They are included here as they did not form part of our original hypotheses, but nonetheless were obvious alternatives that we wished to rule out explicitly.

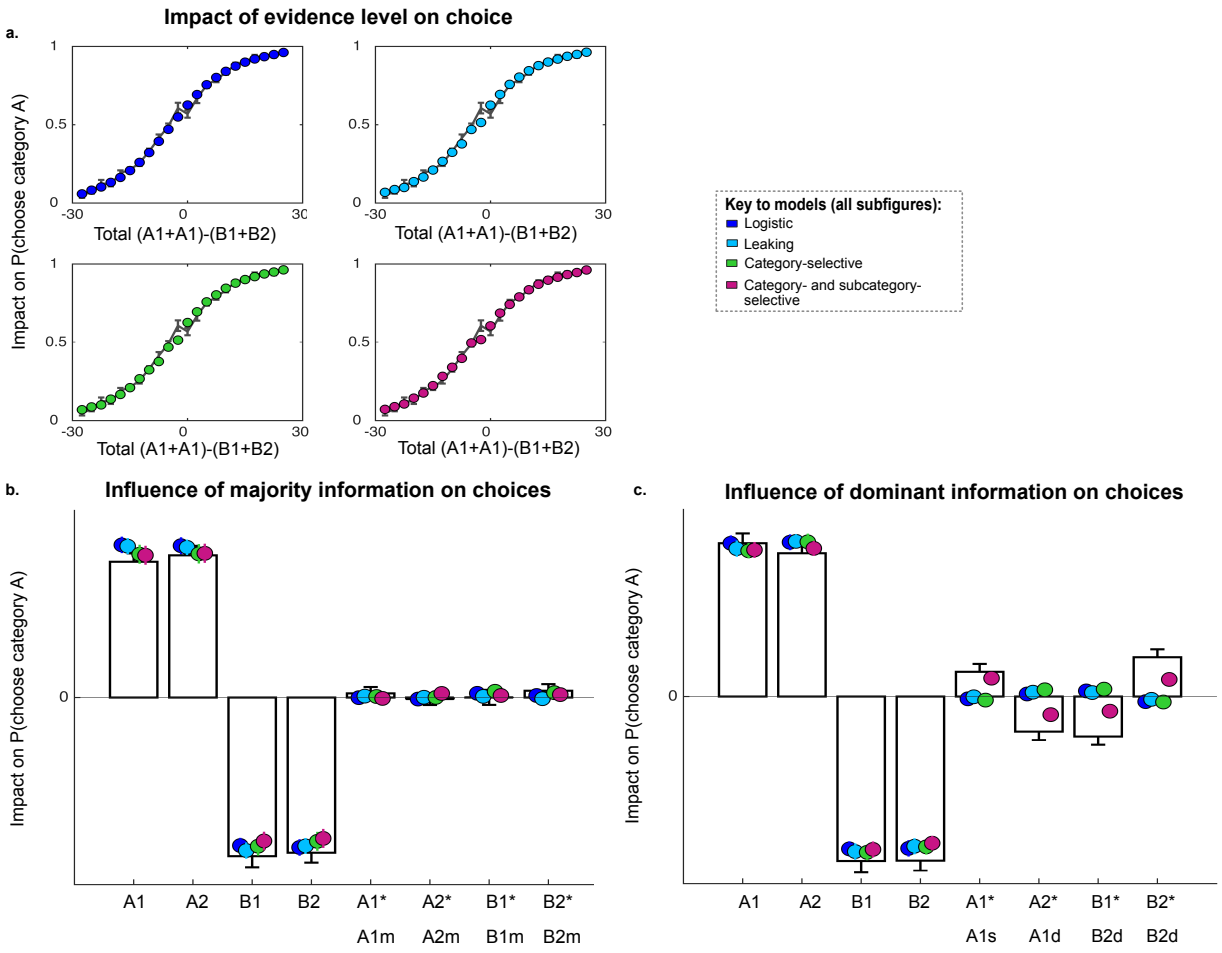

Figure 3.S1. Performance of alternative models, and control analysis. (A). Psychometric curves (i.e. P(choose A) as a function of the optimal decision variable) for the logistic (blue), leaky (cyan), category-selective (green) and doubly-selective (pink) models, in comparison to human data (lines). (B). Impact (beta coefficient) of humans (bars) and the models (coloured dots) choosing 'category A' of information in the four subcategories (A1, A2, B1 and B2; first four bars/dots), and the interaction between this quantity and whether or not the information formed the majority category (final four bars/dots). i_maj was 1 if the relevant subcategory was in the majority category, and -1 otherwise. (C). Impact (beta coefficient) of humans (bars) and the models (coloured dots) choosing 'category A' of information in the four subcategories (A1, A2, B1 and B2; first four bars/dots), and the interaction between this quantity and whether or not the information was in one of the two dominant subcategories (final four bars/dots). Only the doubly-selective model was able to capture the effect seen in human behaviour whereby the information had an enhanced impact on choice if it was in a dominant subcategory.

4. Human optional stopping in a heteroscedastic world

In the previous two chapters, we have seen how humans make good choices both under conditions of low late noise (chapter two, which used a simple task) and higher late noise (chapter three, which used a more complex task). However, a decision can only be 'good' if it is made at the appropriate time: when making decisions, agents must trade off the benefits of information harvesting against the opportunity cost of prolonged deliberation. Deciding when to stop accumulating information and commit to a choice is challenging in natural environments, where the reliability of decision-relevant information may itself vary unpredictably over time (variable variance or "heteroscedasticity"). We asked humans to perform a categorisation task in which discrete, continuously-valued samples (oriented gratings) arrived in series until the observer made a choice at their own pace. Human behaviour was best described by a model that adaptively weighted sensory signals by their inverse prediction error, and integrated the resulting quantities to a collapsing decision threshold. The model mimicked the output of a Bayesian model that computed the full posterior probability of a correct response, and successfully predicted adaptive weighting of decision information in neural signals. Adaptive weighting of decision information may have evolved to promote timely choices in heteroscedastic environments.

In this chapter, we move away from the urn and balls paradigm employed in the previous two chapters. I refer to the urn and balls task as a *comparison* task: the choice being made is "A or B", and information in favour of both categories is presented to participants simultaneously. Whilst in the tasks described in the previous two chapters, information in favour of A will be negatively correlated with information in favour of B (i.e., the more "A" information there is on screen, the less "B" information there is likely to be), it is not perfectly predictive: looking at information from only one category would not allow you to make a fully informed inference. An alternative type of task is a *categorisation* task. The choice is still the same – A or B – but in *categorisation* tasks, only one piece of information is available at a time, and the amount by which this information favours "A" is perfectly

negatively correlated with the extent to which it favours B – i.e. if the information strongly favours A, it strongly disfavours B by the same extent.

In this chapter, I make use of a categorisation task. This is largely in order to make the work more directly comparable to previous work on this topic. However, the two are not unrelated, and I will expand in chapter 5 on the overall insights these two different approaches can give us.

4.1 Introduction

Time is of the essence for an agent who wishes to make good decisions. Longer deliberation promotes decision accuracy, by allowing precise estimates of noisy sensory variables to be formed over time. However, deliberation incurs an opportunity cost, postponing the receipt of the positive reinforcement signals that accompany good choices. Excessive deliberation can thus be pernicious, as exemplified by Buridan's classic fable in which a donkey who is both hungry and thirsty expires whilst choosing between proffered food and water. Successful decision policies must thus strike a delicate balance between the acquisition of sufficient choice-relevant information and the timely harvesting of rewards. Understanding how humans and other animals negotiate this problem has long been a central concern in psychology, neuroscience and behavioural ecology.

Over the past half century, a normative framework has been developed to understand how an agent should decide to stop sampling information and commit to a binary choice (Bogacz et al., 2006; Drugowitsch, DeAngelis, Klier, Angelaki, & Pouget, 2014; Drugowitsch et al., 2012; Gold & Shadlen, 2001; Malhotra et al., 2017; Moran, 2015; Wald & Wolfowitz, 1949). The genealogy of many current theories can be traced back to the Sequential Probability Ratio Test (SPRT), which proposes that agents aggregate the (log) likelihood ratio of evidence up to a fixed threshold, at which point a response is initiated (Gold & Shadlen, 2002; Wald & Wolfowitz, 1949). This policy allows decision commitments to be made at a desired level of certainty about the response , and motivates the popular "drift-diffusion" model (DDM), under which cumulative estimates of noisy inputs drift towards a flat decision bound (Ratcliff & McKoon, 2008). The drift diffusion model has the advantage of approximating the behaviour of the SPRT without the unrealistic requirement that humans know in advance the full distribution from which the evidence is being sampled (Summerfield & Tsetsos, 2012). In the psychophysical literature, the DDM has drawn strong support from the finding that after fitting to response times, the parameters obtained are sufficient to predict the psychometric function that relates decision accuracy to the strength of sensory signals (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008).

For optional stopping problems, models based on a fixed bound are only optimal when the observer can infer the appropriate trial-wise mapping between sensory signals and the probability of a correct

response. This allows the bound height to be flexibly adjusted to optimise speed against accuracy, and thus maximise rates of return (Bogacz, 2007). However, in many natural environments, evidence may vary unpredictably in both its strength (i.e. mean) and reliability (i.e. variance), so that the relevant mapping (likelihood function) is a priori unknown on each trial. In other words, the world is "heteroscedastic". For example, the evidence presented in a law court may be strongly indicative of crime culpability or merely circumstantial (evidence strength), but the witness reporting the evidence might be trustworthy or untrustworthy (evidence reliability). Similarly, economic prospects, such as an investment opportunity, may vary unpredictably in terms of both the mean and variance of a distribution of possible outcomes, i.e. in both their expected value (strength) and risk (reliability). In heteroscedastic settings, reward-maximising solutions to the optional stopping problem can be computed using tools from dynamic programming, but these calculations incur prohibitive computational costs and may be biologically unfeasible for fast sensorimotor judgments (Drugowitsch et al., 2014; Drugowitsch et al., 2012; Frazier & Yu, 2008; Malhotra et al., 2017). The assumptions under which evidence likelihood is computed, and the form of the bound that prompts decision commitment, thus remain key points of controversy in psychology and neuroscience (Boehm, Hawkins, Brown, van Rijn, & Wagenmakers, 2016; Hawkins et al., 2015; Malhotra et al., 2017; Thura et al., 2012).

Here, we investigated the neural and computational mechanisms underlying human optional stopping in a heteroscedastic environment. We used an expanded judgment task, in which humans viewed streams of successive oriented gratings, and indicated in their own time whether grating tilts were drawn from a distribution whose mean fell clockwise (CW) or anticlockwise (ACW) to a reference angle. Our approach differs from those taken previously in three important ways. Firstly, decision information on each trial consisted of discrete, continuously-valued samples of evidence (i.e. grating tilts). Unlike many previous psychophysical studies, this discrete-sample approach allowed us to quantify precisely the available decision-relevant information on each sample, and to model decisions and their latencies under a family of computational models that make qualitatively distinguishable hypotheses. Secondly, we manipulated two distinct sources of uncertainty – the mean and the variance of the generative distribution over tilt – neither of which was signalled to participants on each trial. In this setting, a non-time-varying bound will fail to maximise reward, and normative solutions based on the SPRT require Bayesian inferences about the most likely generative statistics in order to compute, rather than mere summation of sensory signals. This thus allowed us to ask directly

whether participants engaged in signal summation (as in the DDM) or whether they accurately computed the posterior likelihood of each sample conditioned on the possible means and variances (as in the SPRT). For example, in a highly variable stream of evidence, a stronger sensory signal (e.g. a grating tilted far from the reference) is less diagnostic of the correct answer than in a less variable stream, and so optimal deliberation times will be longer for more variable evidence. Thirdly, we allowed participants to draw samples at their own pace. This additionally allowed us to measure, using a single-trial analysis approach, how neural signals accompanying each sample covaried with both time and evidence during the formation of decisions. Specifically, we tested for the emergence of an "urgency" signal that grew with the occurrence of each discrete sample, up to the point at which a choice was made.

In this task, we found that human decisions and stopping times were captured by models in which evidence was accumulated to a bound that collapsed over time, but not by fixed-bound models. Over three experiments, collapsing-bound models fit the data qualitatively and quantitatively better than fixed-bound models on every metric tested. This behavioural finding was supported by the observation of an evidence-independent neural marker, observed over the parietal and central cortices, that built up over discrete samples towards a response. This signal is indicative of a time-varying "urgency" signal that drove decisions to a close, even in the absence of an explicit deadline (Churchland et al., 2008; Drugowitsch et al., 2012; Hanks, Kiani, & Shadlen, 2014; Murphy, Boonstra, & Nieuwenhuis, 2016; Thura et al., 2012). This provides an unusually clear demonstration of the mechanisms by which human decisions about sensory signals are drawn to a close.

We also found that models based on mere summation of sensory signals (i.e. based on the DDM) fared more poorly as explanations of human behaviour than those that computed posterior probabilities of the mean and variance of the generative distributions (i.e. based on the SPRT). We leveraged both behavioural and neural data to ask how such a computation might be implemented. Inspired by previous work involving categorisation of fixed-length sequences of samples, we show that the behaviour of the SPRT can be approximated by a model that accumulates sensory signals into a decision variable in a biased fashion, with reduced weight given to sensory samples that are conditionally surprising, implemented via an adaptive gain control process (Cheadle et al., 2014; Louie, Khaw, & Glimcher, 2013; Niwa & Ditterich, 2008). The decision variable computed by this adaptive weighting model approximated the SPRT, and allowed it to mimic both humans and the SPRT

on a range of qualitative and quantitative metrics. Moreover, the adaptive weighting model successfully predicted that neural encoding of decision information in EEG signals is attenuated on samples which are inconsistent with current beliefs.

These findings suggest that human optional stopping in a heteroscedastic world is dictated by a frugal decision policy, that is based on summation of sensory signals, but adaptively down-weights decision information by the surprise it engenders (Summerfield & Tsetsos, 2015). This may be implemented via a normalisation mechanism that allows decision information to be efficiently coded neurally (Carandini & Heeger, 2012). Adaptive gain control in optional stopping tasks approximates the decision policy of a model, based on the SPRT, that optimally integrates posterior likelihood ratios to a collapsing bound.

4.2 Methods

Eighty-five participants, recruited from the University of Oxford and the surrounding area (experiments 1 and 3, EEG version), and from University College London (experiments 2 and 3, behaviour-only version), took part in three separate iterations of the experiment (n=15, n=37 and n=33 respectively). All participants reported normal or corrected-to-normal vision, and reported no history of neurological or psychiatric disorders. Consent was given in accordance with local ethical guidelines. Participants were either paid for their participation, or completed the task as part of an undergraduate research project (table 1). The data from 10 participants in experiment 2 were excluded due to task non-compliance (a modal reaction time of 2 samples or fewer (n=9), or strong bimodal response time distribution at 2 and 3 samples (n=1). The latter was removed post hoc due to the unusual response pattern, but prior to any statistical analysis). In experiment three, 17 of the 33 participants participated in the EEG experiment, of whom two were excluded due to excessive movement and/or recording artefacts. Of the thirteen additional datasets from experiment three, which were collected as part of an undergraduate project, 5 were excluded due to poor performance (range 49-57% correct in contrast to 68-80%). Excluded participants were removed from all further analyses, leaving n=15, n=27 and n=23 for experiments 1, 2 and 3 respectively (total n=65).

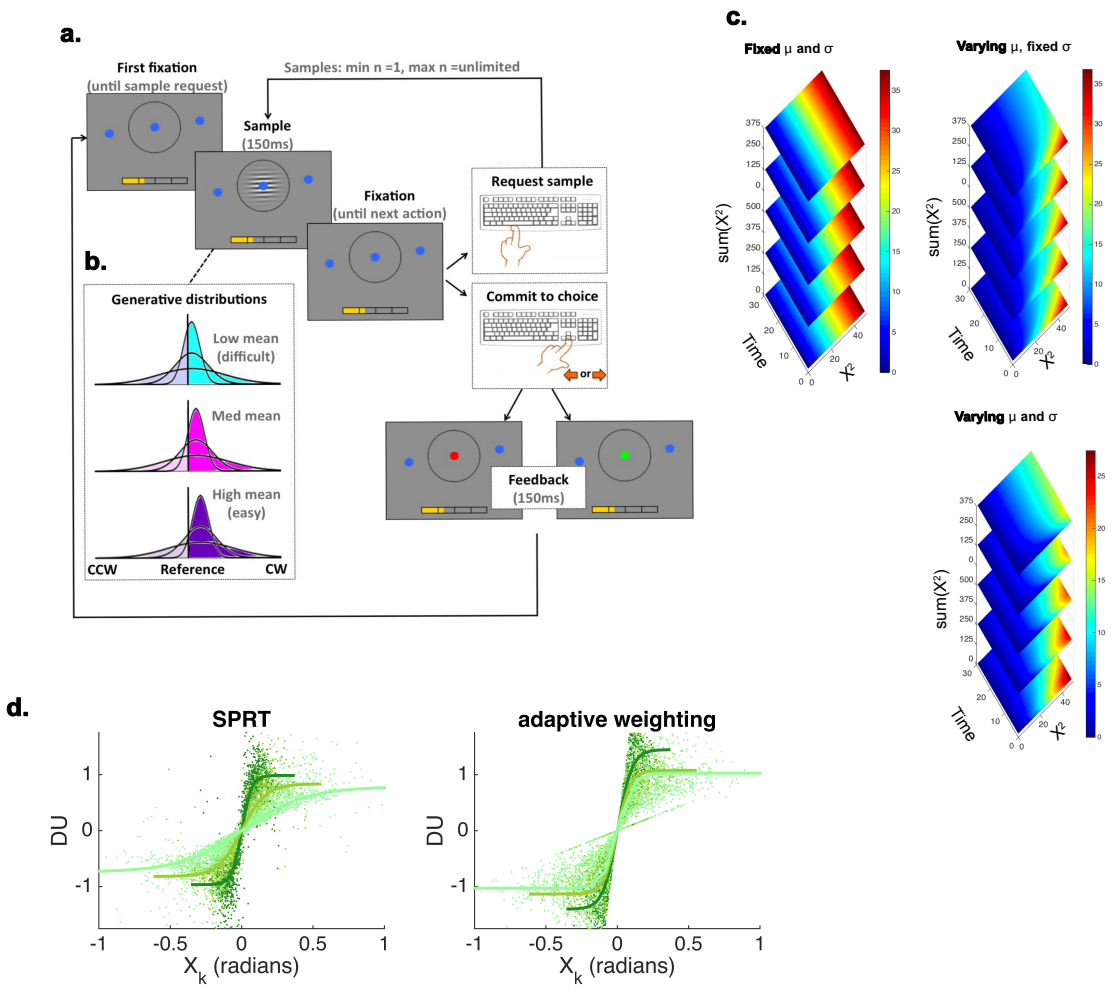

Figure 4.1. Task and simulations. A. Schematic of an example trial. Participants first viewed a circular placeholder with a central blue fixation point, and two further blue dots indicating the reference orientation (frame 1). They pressed the space bar to request a sample, which appeared with the place holder for a 250ms (frame 2), followed by a further fixation screen (frame 3). After each sample, they could either press again to request another sample, in which case they returned to frame 2 (with a different grating drawn from the same distribution), or make a choice by pressing one of the two arrow keys (frames 4/5), upon which they received feedback in the form of the fixation dot turning green (correct) or red (error). Cumulative winnings were shown in a yellow bar in the lower part of the screen. B. The generative distributions for low mean (cyan), medium mean (pink) and high mean (purple), for trials on which the mean was 'clockwise'. The darker colour on each plot reflects the probability of the stimulus having an orientation clockwise of the reference, and the paler colour reflects the probability of the stimulus having a anticlockwise orientation. The distributions for trials on which the correct answer was anticlockwise were identical, but with the reference line on the opposite side of the distribution means. C. The dependence of p(correct) on the sum of evidence (X), the number of samples elapsed, and the sum of squared evidence (X^2) in various environments. p(correct) is shown as the absolute of the likelihood ratio (colour scale). Top left panel: where the mean and variance of the generative distributions is known to the observer. Note that here, sum(X)

is a sufficient statistic for computing p(correct). Top right panel: where the variance is known to the observer, but the mean is not. Here, p(correct) depends on both sum(X) and time, as previously described (Kiani & Shadlen, 2009). Bottom: where neither the mean nor the variance of the generative distribution is known to the observer. This is the case in the experiments conducted here. Time and sum(X) are no longer sufficient statistics for computing p(correct). However, it can be computed if some estimate of the dispersion of the evidence, such as sum(X^2), is additionally computed. D. Mapping function from X (tilt) to DU (decision update) for the SPRT and adaptive gain models. Although the functions are different, they both involve steeper slopes for low-variance distributions (darker colour) than high-variance distributions (lighter colour). Dots are individual trials; lines are best cubic fits. Note that for the DDM, all points would lie on a straight line.

Task design and stimuli

In all versions of this self-paced categorisation task, participants determined on each trial whether the orientation of a stream of stimuli (Gabor patches) was on average clockwise or anticlockwise with respect to a reference (Fig 1a). The stimuli on each trial were drawn independently from a distribution of angular orientation whose mean was ±2, 4 or 6 degrees relative to a reference (±0, 2 or 4 degrees on experiment 1, with 66% of trials in the 0 mean condition leading to 'correct' feedback, on a pseudorandom basis), and whose variance was 4, 8 or 16 degrees. Mean and variance were generated independently and randomly on each trial. Participants viewed as many stimuli as they desired on each trial, pressing the spacebar (experiments 2 and 3) to request each stimulus, until they felt confident enough to commit to a response using the right hand. In experiment 1, stimuli were presented automatically at a rate of 2Hz rather than being actively requested. Participants completed 648 and 540 trials in experiments 1 and 2 respectively. In experiment 3, participants performed the task for exactly 1 hour, with no upper or lower limit on the number of trials they would complete.

The experiment took place in a quiet darkened room with participants' faces approximately 70cm from the presentation computer screen. The visual stimuli were presented using the Psychophysics Toolbox running in MATLAB. Before the experiment began, the instructions were explained clearly and participants had the opportunity to practise the task and ask questions. Each experiment was divided into several blocks whose lengths were determined either by a time or trial limit (table 1). At all times, the display showed a grey screen with a central hollow black circular aperture (figure 1a), flanked by two reference dots (diameter 5 pixels) that would bisect the centre of the circle if joined with a straight line. The location of the reference dots thus determined the reference angle against

which participants compared all incoming stimuli. This location changed randomly block by block, and was drawn from the full range of possible angles (0-180 degrees), meaning that the reference could represent any possible angle within the circle. An additional dot was present in the centre of the circular aperture for clarity (experiment 3 only).

Each trial was initiated when the participants requested the first stimulus by pressing the spacebar (experiments 2 and 3); in experiment 1 the stream of stimuli onset automatically. Each stimulus consisted of an oriented Gabor patch that was displayed for 150ms within the circular aperture. The orientation of the Gabor patch was determined by the mean and the variance of the distribution relative to the reference from which it was drawn (see above). Because the Gabor patches were drawn from a distribution, it was often the case that despite the true mean being 'clockwise', several samples (stimuli) could appear to favour anticlockwise. Thus it was necessary to take several samples in order to ascertain the general tendency (see Figure S1a for p(correct) for a perfect integrator per condition across a trial). Participants continued to sample by pressing the spacebar for each new stimulus (or continuing to view the automatic stream in experiment 1) and committed to a choice (clockwise or anticlockwise) when they felt confident. Visual feedback was given in the form of a green or red central dot (5 pixel diameter) for correct and incorrect responses respectively. At the bottom of the screen was a 'reward bar', which was partially full at the onset of the experiment, and which corresponded to the amount of bonus money a participant could receive. At the same time as the visual feedback, the bar increased by a given increment if participants were correct and decreased by a higher increment if they were incorrect (see table 1 below). This asymmetric payoff structure was necessary to ensure that the strategy to maximise reward rate was not simply to guess the answer after viewing just one stimulus. In experiments 2 and 3 there was a short delay after viewing the stimulus before participants were able to request another stimulus or commit to a choice. This ensured that stimuli couldn't be requested by accident, and gave time (in EEG) for motor signals to abate. Between trials there was a slightly longer delay during which participants could not initiate a new trial; in experiment 2 the aperture and reference dots disappeared during this time, in experiment 3 the feedback red or green remained on until participants were able to begin a new trial. In experiments 1 and 2, the maximum number of stimuli a participant could view per trial was 20; if they reached 20 without making a choice then "incorrect" feedback was shown and a new trial began. In experiment 3, the number of stimuli a participant could view on any trial was 100, this number was purposely chosen to be far higher than we (correctly) assumed any participant would go. However,

had any participant requested 100 samples, pressing the spacebar would no longer have had an effect and they would have been obliged to commit to a choice in order to move on to the next trial.

Table 4.1: Summary of minor methodological differences between the three versions of the experiment

	Experiment 1	**Experiment 2**	**Experiment 3**
Aperture/gabor diameter	100px	100px	190px
Number of blocks	6 x 108 trials	5 x 108 trials	4 x 15 minutes
Experiment length	648 trials	540 trials	1 hour
Reward/penalty amount	+1 reward -1 penalty	+1 reward -3 penalty	+2 reward -4 penalty
Reward bar length at experiment onset	Empty	Half full (150 px), reset every block	One third full (100 px), did not reset
Min delay between stimuli	50ms	150ms	Jittered 0.2-0.3 seconds
Min delay between trials	1000ms	150ms	Jittered 1-2 seconds
Bonus structure	Could earn £1 per block, p(reward) corresponded to how full reward bar was	Could earn £2 per block, p(reward) corresponded to how full reward bar was	£2.50 per quarter of max length of reward bar at the end of experiment (e.g. if it was half full, bonus was £5)
Payment	£8 plus bonus	£5 plus bonus	£15 plus bonus (EEG), voluntary (part of course) (behaviour only)
Max stimuli viewable per trial	20	20	Participants given no limit
Response buttons for choice	R/L mouse button for CW/ACW	R/L mouse button for CW/ACW	R/L arrow key for CW/ACW

EEG acquisition and preprocessing

EEG signals from 64 Ag-AgCl electrodes, plus four additional electrodes used in bipolar montage as horizontal and vertical EOGs, and two electrodes used as references at the mastoids, were recorded using a Neuroscan EEG system with NuAmps digital amplifiers. The EEG signals were recorded at a sampling rate of 1 kHz and were high pass filtered online at 0.1Hz. The impedances of all functioning electrodes were below 10 kΩ at recording onset.

All preprocessing was completed using MATLAB custom scripts and the EEGLAB toolbox (Delorme & Makeig, 2004). First, data were downsampled to 250Hz, band pass filtered between 1 and 40Hz, then split into epochs of -1 to 1 second from each stimulus (sample) onset. It was necessary to epoch sample-wise rather than trial-wise due to the uneven number of samples taken per trial, and therefore widely varying trial lengths. The data from all channels were then visually inspected to remove epochs containing high frequency noise typical of muscle artefacts or electrical surges. Rejected samples formed no further part of the EEG analyses. Data from consistently bad channels (observed either during recording, during visualisation of the data, or both) were interpolated spherically (range of channels interpolated per participant: 0-5, average 2.9 channels interpolated per subject). The data were then re-referenced to the average signal. An extended independent component analysis (ICA) was then conducted, and components with activity typical of blinks or electrical noise were rejected. Finally, data were baselined relative to the first 250ms in each epoch (-1000 to -750ms preceding sample onset).

Models

SPRT Model

The sequential probability ratio test (SPRT) specifies the minimum decision time required in order to achieve a given level of accuracy. To achieve this, an estimate of p(correct) is required. This can be computed using a sequential Bayesian approach that infers, after the occurrence of each stimulus, the posterior probability distribution over an exhaustive space of values for μ and σ, and then marginalises over this distribution to compute a (log) posterior ratio for CW v ACW on each sample k, as follows:

$$DU_k^{SPRT} = log\left(\frac{p\left(X_k \sim N\left(\mu \in [\,1{:}\,12], \sigma \in [1{:}\,32]\,\right)\right)}{p\left(X_k \sim N\left(\mu \in [-1{:}\,-12\,], \sigma \in [1{:}\,32]\,\right)\right)}\right) + \varepsilon$$

Equation 4.1

Where DU_k denotes the decision update on sample k, and in all models:

$$\varepsilon \sim N(0, \varsigma)$$

Where ς was a parameter that varied freely from 0.5 to 20 (0.1 to 20 in experiment 1). Given that the mean and variance distributions were unknown to participants, we initiated the SPRT model on each trial with a flat prior across distributions with 24 possible values of mean (±1-12) and 32 possible values of variance (1-32). These values were chosen to have a maximum value that was twice the highest possible generative value.

Drift Diffusion Model

The drift diffusion model (DDM) utilises a less computationally expensive decision update, simply integrating the evidence on each sample (in this case, evidence (X_k) is the angular disparity between the sample and the reference):

$$DU_k^{DDM} = X_k + \epsilon$$

Equation 4.2

Adaptive gain model

Finally, we considered a model that also bases its decision update on the information provided by each sample, but weights this information according to a prediction error term such that highly deviant stimuli are incorporated with lower gain:

$$DU_k^{AG} = \frac{X_k}{c + \left|X_k - \hat{X}_k\right|} + \epsilon$$

Equation 4.3

where $\hat{X}$ is the expected value of X on sample k, given by the mean tilt of all stimuli up to $k-1$, and c is a small regularisation constant (here, 0.05).

Stopping rules

Two versions of each of the three models were defined: one that incorporated a fixed bound, and one that incorporated a collapsing bound. For both versions, decisions were triggered when the model stopping rule (SR) was true:

$$SR_k^{model} = \left(\sum_{j=1}^{k} DU_{model,j}\right) \cdot \gamma^{k-j} \geq \tau - (\lambda k)$$

Where k denotes the current sample within a trial, and γ reflects a leak parameter (akin to the gradual decay of information in memory), freely varying between 0.8 and 1 in the fixed bound models, and set to 1 in the collapsing bound models. The threshold parameter τ also varied freely (range 1-20). Finally, λ denotes an urgency parameter, free to vary between 2 and 10 for the collapsing bound models and set to zero for the fixed bound models.

Free parameter estimation

The best fitting parameter values – noise, threshold, and leak or urgency – were determined via an exhaustive search. The parameter values that gave rise to the lowest mean square error between the model and human reaction time distributions were selected. Model choices were generated on the basis of the sign of the model decision variable at the point that the decision was triggered.

Lookahead models

In addition to our models of interest, we considered two "lookahead" models that, in a manner reminiscent of dynamic programming, 'simulated' future timesteps in order to respond at the point at which expected value was highest. However, these models provided a poor fit to the data (Supplementary methods, Figure 4.S1b and c).

Behavioural analyses

We compared human and model performance in a number of ways. First, we plotted stopping time histograms separately for correct and error trials for the humans and models across 10 (experiments 1 and 2) or 25 (experiment 3) evenly spaced bins. Secondly, we compared condition-wise stopping times (collapsed across sign of mean at equivalent levels). In the low mean condition, variance level makes very little difference to the conflict in the evidence stream: close to half of the stimuli will be drawn from either category, irrespective of the variance level (this is visually evident from the plots of the generative distributions in figure 4.1b). Thus we expected a relative increase in reaction time according to variance level only in higher mean conditions, which we tested using a two way ANOVA (3 levels of mean by 3 levels of variance) on the average stopping times (STs, i.e. the number of samples viewed before commitment) for each condition.

Many previous studies have investigated decision latency by comparing distributions of stopping times. The nature of our design, in which we knew the precise quantities of all choice-relevant information at each sample, allowed us to employ a novel analysis technique: we took the quantities of interest on each sample and used them to predict whether or not a commitment would be made after that sample. In other words, we ascertained which quantities may be the most predictive of decision termination. To this end, we regressed the cumulative decision information, time elapsed, and the squared decision information (as a proxy to variance) on a given sample, onto a categorical output variable that was 1 if the participant committed to a choice following that sample (stimulus), and 0 otherwise. We then repeated this procedure with the model stopping times for comparison. All collapsing bound models predict an influence of time elapsed, the DDM and the AG predict a positive

influence of the absolute decision evidence. With respect to variance, the DDM makes the opposite prediction from the other two models: according to the DDM, extreme samples (i.e. those particularly found in more variable streams) will speed decisions, as the large jumps in evidence make it more likely that the decision variable crosses a bound earlier. However, for the Bayesian SPRT, extreme values are probabilistically less likely, and are therefore associated with a weaker prior and lower integration gain, making them less likely to cause a threshold to be breached and a decision made. Thus the SPRT predicts that higher variance will slow decisions. The AG model also predicts a downweighting of extreme information, due to its tendency to lead to a larger disparity (prediction error) with the running evidence, thus also leading to lower gain and slower decisions.

A model that simply sums all evidence, such as the DDM, expects a linear relationship between the amount of evidence (angular disparity) each stimulus contains and that sample's influence on choice, with no difference in this relationship as a function of the variance of the generative distribution. The SPRT and the AG models on the other hand both assume a nonlinear relationship that varies across variance levels (see Figure 1d). This is because both models predict lower gain associated with more variable stimuli. To ascertain which of these patterns (linear versus nonlinear) more closely captured human performance, we conducted a robust regression, binning each sample according to its size (i.e. its angular disparity from the reference) and using this to predict its impact on choice, separately for each condition of variance, separately for human and model choices.

Model comparison

All formal model comparisons were conducted using the VBA toolbox (Daunizeau, Adam, & Rigoux, 2014). All models had an identical number of free parameters and thus no adjustments were required to the scores entered into these comparisons.

EEG analyses: encoding

In many cases, different mechanisms (such as those proposed by each model) can lead to similar behaviour choices. Turning to the neural recordings taken from a subset of participants in experiment three allowed us to ascertain directly which underlying mechanism best accounts for neural data.

The adaptive gain model calculates its decision update by summing the evidence provided on each sample after weighting it according to the prediction error (see equation 4.3). Thus it predicts signals corresponding to the momentary level of evidence (the angular disparity between the stimulus and the reference), the prediction error, and the interaction between the two.

The SPRT model also predicts a representation of the level of evidence on the sample, as the quantity is central to its decision update process. Over and above this, the SPRT model would predict a neural representation of the log posterior ratio of evidence, i.e. the quantity the model actually integrates (see equation 4.1).

To ascertain which (if either) of these predictions are supported by our data, rather than calculating event related potentials (ERPs), we took parametric predictors within a general linear regression model to predict sample to sample fluctuations in the average EEG activity across frontal (F1, Fz, F2, FC1, FCz and FC2) and parietal (P5 P3 P1 Pz P2 P4 and P6) electrodes. We evaluated the following regression model:

$$EEG_t = \beta_0 + \beta_1 \cdot |X| + \beta_2 \cdot |DU_{SPRT}| + \beta_3 |X - X^{px}|$$

$$+ \beta_4 (|\bar{X} - X^{px}| * |X|) + \beta_5 \cdot \log(N') + \beta_6 \cdot |X^{px}| + \beta_7 \cdot |X^{nx}| + \beta_8 \cdot F + \beta_9 \cdot L$$

Equation 4.4.

$|X|$ represents the (absolute) momentary decision information conveyed by the relevant sample. $|DU_{SPRT}|$ reflects the log posterior ratio of evidence, i.e. the momentary quantity predicted by the

SRPT model (equation 1). The prediction error term $|X - X^{px}|$ denotes the absolute angular difference between the current stimulus and the mean of all previous stimuli in the current trial up to that point, thus generally yielding higher values for samples from more variable streams. The adaptive gain model predicts an interaction between this prediction error term and the level of evidence X, reflected by β_4, as information that is more discrepant should generate a lower update. The final five terms ($\beta_5 - \beta_9$) are nuisance predictors that signal decision information on the previous ($|X^{px}|$) and subsequent ($|X^{nx}|$) samples, and variance associated with the first sample (F) and last sample (L; i.e. the sample that prompted commitment) in the trial. Finally, N', encodes proximity to the eventual choice, equivalent to the "response-locked" regressor sometimes employed in psychophysical studies. N' starts strong and gradually weaken as the trial progresses, akin to an adaptation signal. The regression was repeated for EEG activity averaged across parietal as well as frontal regions.

We regressed all 9 predictors against the single trial EEG activity from the corresponding trial, point by point over 500 evenly spaced datapoints that spanned from -1s to 1s from sample onset. The resulting encoding curves reflect parameter estimates: they show the slope of the best fitting regression line at each time point, thus indicating how strongly the EEG signal varies with the model-predicted quantities. Given the entirely self-paced nature of this experiment, epochs from adjacent samples sometimes had a degree of overlap. However we believe that the responses to each individual sample can be sufficiently disambiguated due to the inclusion of pertinent information from the surrounding samples as nuisance regressors, as described above. Furthermore, information from one sample to the next was decorrelated particularly due to the method used, which involved ignoring information about trial membership: all samples from all trials were regressed sequentially against the EEG signal, meaning that samples from a different generative distribution followed on the heels of those that had formed part of a different trial with different underlying structure. The methods we used for disambiguation are akin to the principles in parametric event related functional neuroimaging designs (Josephs, Turner and Friston 1997). Entering all of the quantities in the same regression ensured that they competed for unshared variance, meaning that we could be confident that any effects were the result of the predictor's unique influence on the EEG signal. The resulting slopes were then averaged across samples, and statistical analyses were conducted at this group level. Significance of beta deviations from 0 was assessed using a nonparametric cluster correction technique, implementing a familywise error (FWE) with an alpha of 0.05.

EEG analyses: Lateralised beta band activity

Previous studies have shown beta band desynchronisation in the build up to motor execution of a decision in the contralateral hemisphere to the hand making the choice. In our experiment, samples were requested with the left hand and a decision commitment was made by the right. Thus we would expect to see desynchronisation in the left hemisphere, i.e. contralateral to the response hand. Furthermore, this signal should be influenced by quantities that are predictive of choice execution. In particular, our paradigm allowed us to determine whether the cumulative decision information and/or time elapsed were more predictive of this lateralised desynchronisation motor preparation signal.

To this end, we computed the interhemispheric difference in neural activity across 10 logarithmically spaced frequency bands (~10-40Hz) at lateral central electrodes, by subtracting the spectral log power of C3+CP3 from C4+CP4. We then regressed the absolute decision information predicted by the models: the cumulate of the angular disparity, the cumulate of the log posterior ratio and the cumulate of the prediction error, as described in the encoding analysis above, as well as the log of the sample number (i.e. time elapsed) against the resulting signal, in order to ascertain whether there was an evidence-independent neural signature of urgency. As a control, we also included the inverse of the urgency signal, i.e. the quantity that started large and decreased over the course of a trial, to control for the large neural adaptation signal typically seen. The resulting parameter estimates were averaged across samples separately for each frequency band before being entered into group-level statistical analyses, in which we used a nonparametric cluster correction technique with familywise error of alpha =0.05 to assess significance.

4.3 Results

In the general form of the decision problem discussed here, observers view successive samples of information, each characterised by a continuously-valued feature. In the experiments described here, feature values denote the tilt of each grating relative to a reference orientation. Observers may view as many samples as desired before indicating their belief that they are drawn from a Gaussian distribution with positive or negative mean, $N(+\mu, \sigma)$ or $N(-\mu, \sigma)$ (i.e. that in the limit, they are on average tilted CW or ACW of the reference angle). In the experiments we describe, we varied μ and σ each at three levels relative to the reference. The task was performed for a fixed time period or number of trials, with real financial incentives for correct and incorrect responses (see **Figure 4.1a** and methods for details).

Models

The probability of a correct response can be computed using a sequential Bayesian approach that infers, after the occurrence of each sample, the posterior probability distribution over an exhaustive space of values for μ and σ, and then marginalises over this distribution to compute a (log) posterior ratio for CW vs ACW. This model is known as the sequential probability ratio test (SPRT). For heteroscedastic inputs, the likelihood function that maps sensory features onto decision values evolves dynamically as beliefs about the most likely values of μ and σ change over the course of the trial. We note that this is not the fully optimal policy, which would require estimation of the expected value of responding now vs. later using dynamic programming; however, models incorporating these calculations offered a poor fit to human data (see **figure 4.S1b** and c).

In a hypothetical case (not characteristic of the experiments here) where μ and σ are known to the observer, this SPRT model is equivalent to one that integrates the sum of evidence, i.e. to the DDM. Under the assumption that observers (erroneously) assume that information is drawn from a distribution with a single value for μ and σ that is fixed over all trials, thus, we considered the DDM here as it is the natural rival to the SPRT for explaining our data. Finally, we proposed a related model that sums the grating angles over time after modulation by a normalisation term that depends on prediction error. This model implements an adaptive weighting process whereby samples that are

inconsistent with the running mean (i.e. surprising samples) are down-weighted relative to those that concur with expectations. We chose to include this model because a very similar account captures qualitative features of human behaviour in tasks involving a fixed number of samples.

In **figure 4.1**, we plot the decision values for the SPRT and adaptive weighting models as a function of X for an example participant dataset, separately for each level of variance of the generative distribution σ (for the DDM, the decision values are simply equal to X_k, and so the points would lie on the identity line). Note that for both the SPRT and adaptive weighting models, the mapping from inputs to decision values varies with σ in a similar fashion. These two models thus approximate each other and should be expected to perform similarly in explaining qualitative aspects of human behavioural data.

Behavioural data and fitting

Using a first passage process – i.e. one that determines that point at which the decision variable first hits the bound on a given trial – we fit the SPRT, DDM and adaptive weighting models to human stopping times on correct and error trials in 3 behavioural experiments (Exp.1, n = 15; Exp.2, n = 24; Exp.3, n = 27; see methods for details). Exp. 2 and 3 differed in minor ways; in Exp. 1 gratings were shown at a rate dictated by the experimenter (4 Hz), rather than being sampled in a self-paced fashion. For all models, we allowed the boundary height and noise term ε to vary as free parameters. For fixed-bound variants, we also included a leak term that captured potential loss of information during integration. For collapsing bound models, we replaced the leak parameter with a term that allowed the threshold height to decline over time by a fixed amount on each step, leading to a linearly collapsing bound. We included both fixed- and collapsing-bound versions of each of the three model variants, yielding 6 in total. In what follows, we compared models via quantitative metrics (i.e. log likelihoods) and random effects analysis (i.e. Bayesian model selection), but we also simulated best-fitting variants of our models to generate predictions about qualitative features of the human data, allowing us to draw strong conclusions about the computations underlying optional stopping based on model falsification (Palminteri et al., 2017).

Stopping time distributions.

We first compare the fits to stopping time distributions for Exp. 3 (the EEG dataset) in **figure 4.2a** (see Figure 4.S1d and e for Exp.1 and Exp.2 respectively). Visual inspection suggests that the collapsing-bound models fit better than the fixed-bound models in all three experiments, and this was confirmed by frequentist comparison of the mean-squared error for every pairwise comparison (Exp.1: all t-values > 6.42; Exp. 2: all t-values >5.18; Exp. 3: all t-values > 2.42; all p-values < 0.001 for experiments 1 and 2, all p<0.05 for experiment 3). For a more formal quantitative analysis we pooled over all three experiments and compared model log likelihoods using Bayesian model selection. This analysis strongly favoured the collapsing bound models (expected frequencies for fixed-bound models were all less than 2%, whereas chance is 16.7% for the 6 models; and 31.4%, 23.6%, and 41.3% respectively for the SPRT, DDM and adaptive weighting models with collapsing bound). Expected frequencies for the collapsing-bound SPRT and normalisation models were above chance ($p < 0.002$ and $p < 0.001$ respectively) but did not exceed chance for the DDM with collapsing bound ($p = 0.08$).

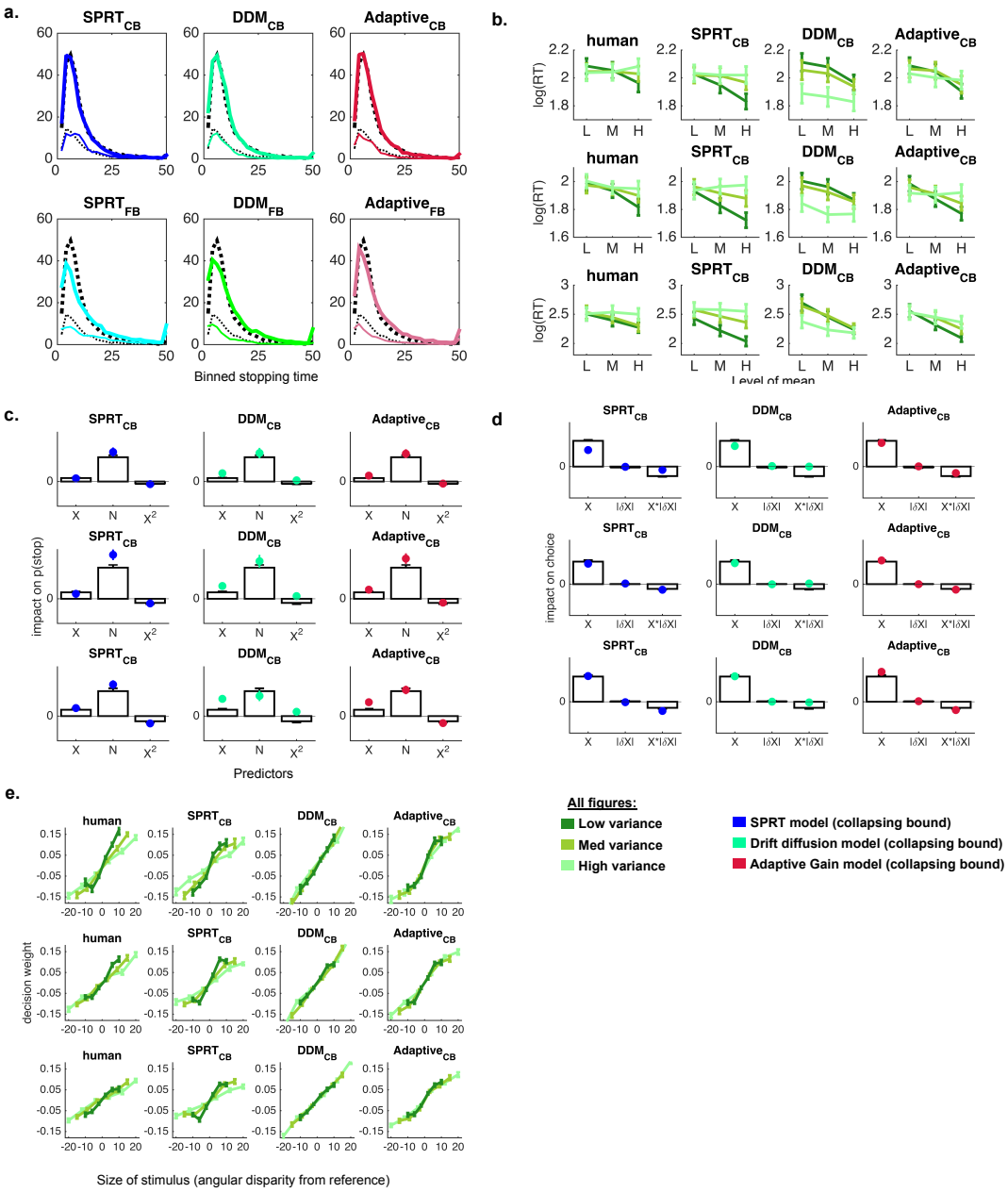

Figure 4.2. Behaviour and modelling. (A): human distributions of stopping times for correct (thicker black line) and error (thinner black) trials, and the corresponding fits of fixed- and collapsing-bound variants of the DDM, SPRT and adaptive weighting models (thick and thin coloured lines). Collapsing bound models fit better. (B): average (log) stopping times (in samples) for trials with low, medium and high mean, i.e. distance to the reference (x-axis) and low, medium and high variance, i.e. distance to the reference (dark, medium and pale green respectively). Leftmost panel: humans; other panels, models. Rows 1-3 are experiments 1-3 respectively. The particular increase in reaction time as a function of variance in the high (easy) mean condition can be explained by the task structure: when the mean is low (hard), the ratio of samples favouring either category (CW v ACW) will be almost 50:50, with only a very small bias towards the generative correct category, irrespective of the variance. However when mean is high, the variance condition has a much larger influence on the ratio

of samples in favour of either category (see figure 1 for an illustration of the generative structure). (C): coefficients from regressions predicting next-sample commitment based on sum(X), samples elapsed (N) and sum(X^2). White bars are humans; each panel shows the fit of a model (blue, green and red dots reflect SPRT, DDM and AG models respectively). Rows 1-3 are experiments 1-3 respectively. The $SPRT_{CB}$ and $Adaptive_{CB}$ models capture the negative influence of sum(X2), but the DDM_{CB} does not. (D): the impact on p(choose clockwise) of sample information (X), the prediction error $|\delta X|$, and the interaction of the latter two. The $Adaptive_{CB}$ model predicts a negative interaction, meaning that the strength of impact of a stimulus on the decision decreases as the prediction error associated with that stimulus increases. (E): subjective versus objective weighting of the decision evidence, separately for the three conditions of variance (low, medium and high in pale, medium and dark green respectively). The impact of more extreme angles, and angles from more variable streams, is lower in the human data. This is captured by the AG and SPRT models, but not the DDM. Cf figure 4.1D, which shows the expected pattern based on simulated data.

The best-fitting variants of each collapsing-bound model were then used (without further assumptions) to make predictions about four qualitative features of the human data: (i) condition-wise mean stopping times; (ii) predictors of next-sample commitment; (iii) adaptive weighting of information by the sample history; and (iv) weighting of sensory features in choice. For these comparisons, we focussed on collapsing-bound models because the fixed-bound models fit the stopping time distributions more poorly. However, for completeness, we present the full results in Fig. 4.S1 (fixed-bound models fit the human data more poorly on every measure considered here).

Condition-wise mean stopping times

First, we compared model predictions about the mean stopping times independently for each level of $|\mu|$ and σ. We show these in **figure 4.2b**. The SPRT and adaptive weighting models all predicted that stopping times should vary as an interaction of by $|\mu|$ and σ, with commitments delayed when the samples were drawn from more variable distributions, but this slowing exacerbated as $|\mu|$ deviates further from zero. This occurs because for both of these models, outlying feature values (i.e. more deviant tilts) are penalised more sharply during conversion to momentary decision signals, leading to shallower slopes of integration where samples are more variable, in particular when μ is further from zero. By contrast, the DDM predicts an inversion of this effect: that that stopping times should be shortest when samples were more variable. This latter prediction follows from the dynamics of bounded integration under the DDM, whereby more variable evidence ensures that the diffusion process occupies a potentially wider range of states, and is thus more likely to contribute to a first passage process, in particular when evidence strength is low (Zylberberg et al., 2016). Human stopping time data displayed in the interaction between $|\mu|$ and σ predicted by the SPRT and adaptive

weighting models in all 3 experiments (Exp.1 $F_{2.6,36.9} = 11.39$, $p < 0.001$; Exp.2: $F_{3.3,85.4} = 5.37$, $p < 0.001$; Exp.3: $F_{3.6,82.1} = 4.41$ $p < 0.004$; ANOVA on log stopping times). Comparing least-squares fits of the models to these data aggregated over experiments, we found that the DDM performed overall worse than the other models (t-test on mean squared error: DDM vs. SPRT $t_{65} = 2.30$, $p < 0.02$; DDM vs. adaptive weighting $t_{65} = 4.39$, $p < 0.004$). This confirms the pattern suggested by the quantitative fits above, and argues against the DDM as a description of human optional stopping in our experiment.

Predictors of next-sample commitment

Next, we considered a second qualitative feature of the data, capitalising on our use of a task that permitted precise quantification of the information provided by each sample (i.e. momentary tilt values). We used probit regression to predict whether participants would decide to commit or defer on a given sample (Malhotra et al., 2017), as a function of three time-varying quantities as predictors: (i) the absolute sum of feature values up to that sample (i.e. the DV computed by a noiseless DDM); (ii) the number of samples viewed thus far; and (iii) the sum of squared feature values. This choice of predictors was motivated by the observation that in the heteroscedastic setting provided by our experiment, these three variables are sufficient statistics for the computation of $p(correct)$ (see **figure 4.1c**). Once again, the DDM makes a prediction that diverges from the other models, in that it proposes that the sum of momentary evidence, rather than the passage of time through the trial (i.e. samples elapsed), should be the strongest predictor of next-sample commitment. The other models predict that samples elapsed is the strongest predictor of next-sample commitment, and that a larger sum of squared evidence has the effect of delaying commitment. This latter pattern was observed in human data, again providing evidence against the DDM as a model of human performance in our task (Figure 4.2c). To assess the significance of this effect, we used coefficients from each model to predict human commitments, and compared the resulting model likelihoods using Bayesian model selection aggregated over the 3 experiments. Comparing among the 3 collapsing bound models, expected model frequencies were 54.6%, 7.7% and 37.8% for the SPRT, DDM and adaptive weighting models respectively, indicating a clear disadvantage for the DDM.

Adaptive weighting by sample history

A third qualitative prediction that can be used to arbitrate among the models concerns how the statistics of tilts observed on each trial determine their weight (or impact) on the categorical choice,

i.e. CW vs. ACW response. It follows from the adaptive weighting mechanism model that sensory samples that are incongruent with those occurring previously will carry less weight in eventual choices, because they lead to larger prediction errors. To quantify this effect in all 3 models, we used a previously-described approach, in which probit regression is used to predict choices (CW vs. ACW) as a function of (i) the sum of sensory evidence (i.e. tilts), (ii) the absolute difference between each grating and its predecessor, and (iii) the interaction between (i) and (ii). In previous studies involving integration of a fixed number of samples, we have observed negative coefficients for (iii), indicative of a downweighting of surprising or inconsistent evidence (a "consistency" bias) (Cheadle et al., 2014). Here, the human consistency bias was significant in all 3 experiments (Exp.1: $t_{14} = 14.04$, $p < 0.001$; Exp. 2: $t_{26}=6.50$, $p<0.001$; Exp. 3: $t_{23}=6.08$, $p<0.001$; and this pattern was replicated by the SPRT and normalisation models, but not by the DDM. (Figure 2d). Quantitative analysis using Bayesian model selection, revealed lower expected frequencies for DDM than other models (2% vs 85% for SPRT and 12% for adaptive weighting model; $p < 0.001$).

Weighting of sensory features in choice

A fourth prediction that differs qualitatively among the models concerns the form of the likelihood function that maps tilt values onto choices for different levels of σ. The approximate expected form of this function for the SPRT and adaptive weighting models is visualised in Figure 2e. In humans, this function was estimated by counting grating tilts that fell within 7 bins of angle with respect to the reference, and entering these tallies into a design matrix that was then used to predict choices via (robust) logistic regression. We conducted this analysis independently for trials defined by each level of σ. The resulting coefficients were plotted for humans, and for the 3 collapsing bound models using an identical analysis of model choices (**Figure 4.2e**). The DDM predicts that this mapping function should be linear, and identical for all 3 variance conditions. The SPRT and normalisation models, by contrast, predict a choice function that is sigmoidal (rather than linear) in form, with outlying tilt values "squashed" or down-weighted. Moreover, they both predict that these sigmoidal choice functions will have a steeper slope for low-variance than high-variance conditions; this effect is more exaggerated for the SPRT model. The human data are shown for comparison in the leftmost column of the figure. We compared slopes of sigmoidal functions fit to these choice functions for each variance conditions separately, finding them to be reliably steeper for the low variance condition in all three experiments (Kruskall-Wallis test, Exp.1: $p < 0.02$, Exp.2: $p < 0.005$, Exp.3: $p < 0.001$). This

finding argues strongly against the DDM, which implements a decision policy that does not account for variability in the generative distribution on each trial.

Together, these behavioural analyses provide a rich picture of the dynamics of human optional stopping in a heteroscedastic environment. We draw three conclusions. Firstly, models based on a fixed bound cannot jointly explain human stopping times and choices. Fixed-bound models fit more poorly than collapsing-bound models on every qualitative and quantitative metric used here, providing a notably clear demonstration that humans do not draw decisions to a close (at least in our task) using a flat decision threshold (Figure 4.2a). Secondly, models that involve unbiased linear accumulation of evidence to a collapsing bound (e.g. based on the DDM) fail to explain the qualitative pattern of stopping times as a function of evidence strength and reliability (Figure 4.2b), to account for how the statistics of stimulation predict next-sample commitment (Figure 4.2c) to capture how the disparity between current and previous sample value influences choices (Figure 4.2d), or to correctly predict the form of the choice function that humans use to map sensory features onto decision evidence (Figure 4.2e). These qualitative observations seem to rule out an unbiased, linear integration process as driving human choices in this task. Thirdly, we report that the slope of the human weighting function (transducer) varied as a function of the dispersion of sensory features in the stimulation sequence. This suggests that humans either inferred the generative statistics of the sequence online during stimulation (SPRT model), or used an approximation that allowed them to behave as if they were doing so (adaptive weighting model).

EEG data

Encoding analyses

In the analyses above, we examined qualitative features of the behavioural data with a view to ruling out various classes of optional stopping model. However, two models – the SPRT and adaptive gain model – were similarly able to predict all the major features of the human data, and on most measures did not differ substantially in their quantitative fits. We thus turned to an analysis of neural signals recorded whilst a subset of participants from Exp.3 performed the task, with a view to arbitrating between these models at the implementational level.

Rather than driving our analysis on the basis of event related potentials (ERPs), we employed a single trial approach (as described previously, (Wyart, de Gardelle, et al., 2012; Wyart et al., 2015), also see methods) that allows us to assess how the strength of the neural signal is influenced by parametric decision-relevant quantities such as the DU predicted by each model, thereby revealing the unique influence of each quantity of interest. To do this, we regressed such quantities against the EEG activity evoked by each sample to derive beta weights (i.e. the slope of the relationship between the quantity of interest and the EEG signal) at successive timepoints following sample onset. The resulting weights were then averaged across the cohort; time periods at which these weights deviated significantly from 0 indicate that the relevant quantity is being encoded with above-chance strength.

Both the SPRT and the adaptive gain models predict an influence of the decision evidence X, as this quantity is used to derive the decision update for both models. However, only the Adaptive Gain model predicts significant encoding of the prediction error term, as well as its interaction with the decision evidence X, as these quantities are used to compute the model's decision update. The DU of the SPRT, on the other hand, which is equivalent to the log posterior ratio of evidence, is only predicted by the SPRT model to influence neural signals.

We regressed these quantities against the neural activity averaged over parietal and frontal electrodes. The results of the analysis are shown in Figure 3a (parietal) and b (frontal). Encoding of |X| exhibited a characteristic negative-positive deflection at ~200/450ms over parietal electrodes (red trace; top row), and the sign reversed pattern in frontal electrodes (bottom row) that closely replicates the encoding of decision information described previously (Wyart, de Gardelle, et al., 2012; Wyart et al., 2015). However, there were also a number of additional findings. In line with the predictions of the adaptive gain model, there is significant encoding of the prediction error signal over parietal and frontal electrodes (again with sign reversed pattern in the two regions) 400ms after stimulus onset. This signals a component of the EEG signal that was modulated when samples viewed were incompatible with the running stream of evidence. Importantly, in both regions there was also a significant interaction between the prediction error and the decision evidence, meaning that the strength of encoding of the decision was modulated by the degree of prediction error, as predicted by the adaptive gain model.

In contrast, parietal regions showed no significant encoding of the LPR, i.e. the quantity accumulated by the SPRT model. The lack of such a signal in parietal regions is particularly pertinent given previous links between this region of cortex and the accumulation of decision-relevant information (e.g. (Hanks, Ditterich, & Shadlen, 2006; Kiani et al., 2008; Kira, Yang, & Shadlen, 2015; Roitman & Shadlen, 2002)). In frontal regions, there was a short period of activity that correlated with the predictions of the LPR model. However, given that it occurred pre-stimulus, it is hard to draw any strong conclusions about its potential role in the evidence accumulation process.

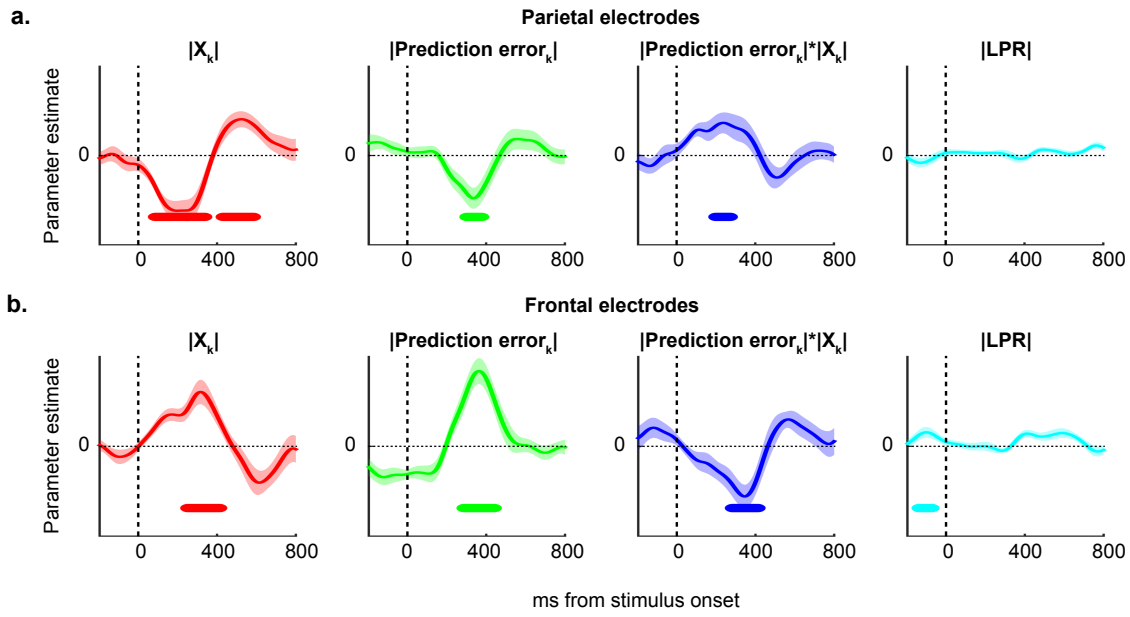

Figure 4.3: Neural encoding of model-predicted information. Curves depicting the correlation between model-predicted quantities and the EEG signal in parietal (A) and frontal (B) regions from -100 to 800ms following each sample. The shaded regions around the solid lines depict SEM. The coloured lines show periods of time during which the signal significantly deviated from 0. The adaptive gain model predicts the encoding of the prediction error signal that is seen here. The SPRT model predicts encoding of the LPR, but no significant encoding is seen in either parietal or frontal regions.

Thus, analysis of the signal over parietal and frontal regions yielded results that clearly supported the notion that at the implementational level, the behaviour predicted by both the adaptive gain and the SPRT models is manifested via the relatively simpler calculations of the adaptive gain model, rather than the computationally expensive mechanism implied by the SPRT.

Lateralised frequency band activity

In behavioural comparisons, collapsing bound models overwhelmingly captured reaction time data more closely than the fixed bound models, implying that a time-sensitive signal that is independent of the cumulative evidence is influential in driving choices to a close. Our next step thus sought to determine whether such an "urgency" signal was present in our data. We motivated the analysis on the basis of previous work indicating that desynchronisation in beta band activity in central regions of the hemisphere contralateral to the hand that executes the motor response is strongly present in the build-up to choice selection. In this task, by design, participants requested samples and committed to choice with the left and right hands respectively. Thus quantities that are strongly linked to choice execution should be strongly predictive of the relative increase of contralateral (left) hemisphere (electrodes C3 and CP3) desynchronisation over ipsilateral (right) hemisphere desynchronisation (electrodes C4 and CP4).

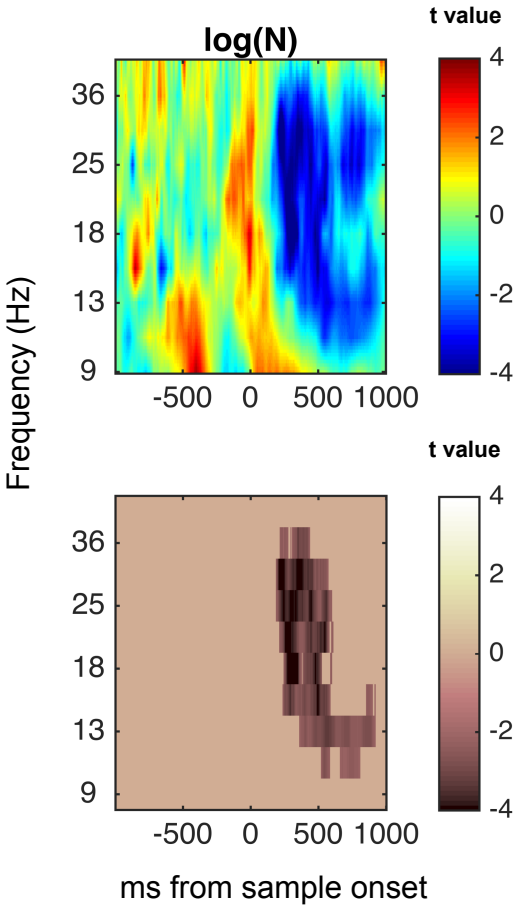

Figure 4.4. T values derived from beta weights from a regression assessing the correlation between the log time elapsed and the interhemispheric difference in the EEG signal from -500 to 1000ms following stimulus onset. Desynchronisation in beta band activity is associated with preparation to execute response; factors that influence this process should therefore cause greater desynchronization. Time elapsed has a highly significant impact on desynchronization; the upper panel reflects overall t values, the lower panel depicts only values which exceed a cluster corrected alpha of 0.05.

We regressed a quantity corresponding to the log time elapsed, alongside the cumulative decision evidence, the cumulative prediction error and the cumulative LPR quantity competitively against the interhemispheric difference of the wavelet transformed central scalp EEG data in 10 logarithmically spaced frequency bands, on each sample from -500 to 500ms following stimulus onset. As an additional nuisance regressor we included the quantity N', which starts large and decreases over the course of a trial (akin to an adaptation signal). We found that, when controlling for the influence of increases of these other decision-relevant factors, there remained a striking and highly significant independent influence of time elapsed, strongly indicative of an 'urgency' signal influencing motor response (**Figure 4.4)**. Interestingly, none of the decision-relevant quantities significantly influenced this lateralised build up.

Summary

Over the course of three behaviour experiments, collapsing bound models were clearly and consistently better at accounting for human data than fixed bound models. Of the three different models, two closely approximated human behaviour: an optimal model ('optimal' in the same sense as the SPRT), and a simpler model that processes discordant evidence with lower gain. Whilst these two models give very similar behaviour results, they implicate very different underlying processing mechanisms, thus neural data were used to arbitrate between the two. The neural signal over parietal and frontal regions correlated with the predictions of the simpler adaptive gain model, rather than the optimal SPRT model.

4.4 Discussion

Lengthy deliberation over a difficult decision is costly: information acquired per unit of time is generally minimal. Fast decisions on the other hand often result in poorer choices. How do agents find the right balance? It has been known for several decades that in order to achieve a given level of accuracy in a free response task, evidence should be accumulated to a fixed (if trial difficulty is constant) or collapsing (if trial difficulty varies) threshold. In these situations, the sum of decision evidence and time elapsed are sufficient statistics to estimate the probability of being correct at any given moment (in the former case). However, decision difficultly is influenced by factors other than just strength of the evidence: choosing the best cycle route to the shop is made harder by the knowledge that the shortest path suffers from constantly fluctuating pedestrian traffic. When evidence variance influences decision difficulty over and above evidence strength, time elapsed and sum of evidence are no longer sufficient statistics to estimate the probability of making a correct choice.

Here, we investigated explicitly what drives agents to commit to a decision when the evidence strength and variance are manipulated orthogonally from trial to trial. We found that an 'adaptive gain' model, which assumes agents sum the evidence provided by each stimulus in a weighted manner that causes inconsistent information to have lower impact on choices, captured human data closely. The qualitative predictions of this model approximated those of a Bayesian SPRT model, which integrates the posterior probability of each stimulus being drawn from the relevant distributions. Despite behavioural similarity, the two models clearly implicate very different underlying mechanisms. Analysis of the concurrent EEG activity allowed us to tease apart the model predictions in a manner that behaviour alone could not, and strongly favoured the adaptive gain model as the best explanation for human evidence integration and decision termination.

The adaptive model implicates a gain control mechanism via which the strength of evidence accumulation is modulated on the basis of the local statistics of the available information, in a manner reminiscent of the adaptation of low level sensory systems to the ambient surroundings (Bartlett, 1965; Carandini & Heeger, 2012; Fairhall, Lewen, Bialek, & de Ruyter Van Steveninck, 2001). The model's ability to explain human data here replicates previous results from a task that employed an

experimenter-paced multi-sample orientation discrimination paradigm: as here, it was found that information was transformed in a nonlinear manner, with enhanced processing of that closest to the running tally of evidence. This down-weighting of statistically deviant information has been observed in a number of other studies: in averaging tasks in which all information is presented simultaneously, outlying information is 'robustly averaged' such that it carries lower eventual weight on choice (de Gardelle & Summerfield, 2011; Li et al., 2017). However, none of these paradigms were paced entirely by the participant – i.e. with each sample arriving at the participant's request until they decided they had seen sufficient information – therefore it was unclear whether, when the viewer dictates the pace, the down-weighting of outlying information is an element of the process that is sacrificed in favour of reaching the bound more quickly, and therefore making a faster decision. Here, we showed clearly that the down-weighting of discrepant information does indeed still feature strongly in self-paced decisions, despite this time cost.

We hope that our data are able to shed further light on the debate of the nature and existence of an 'urgency' signal. Behaviourally, we found that collapsing (rather than fixed) bound models fit our data better in every scenario investigated, supporting the idea that an evidence-independent time-dependent signal does indeed influence choices. However in many other situations, static bounds have been shown to capture evidence better (Hawkins et al., 2015). What may be the cause of these disparate findings? One explanation is trial difficulty: a choice that remains challenging even after a long time has passed is indicative of a low signal to noise environment, meaning that harvesting additional information will have little benefit and it is more beneficial to decide randomly and move on to a more fruitful choice (Juni, Gureckis, & Maloney, 2016). Collapsing bounds are one way to prompt such choices. If the decision is sufficiently easy, there is no need for an additional signal, as bounds will be reached and decisions triggered on the basis of the strength of the evidence alone. Neurally, we found striking evidence of an 'urgency' signal in the frequency band decomposed signal that was independent of the evidence in the trial. We are not the first to show neural evidence for urgency: for example, LIP activity apparently ramps to a bound even in a zero coherence condition, which may be attributed to an urgency effect (Churchland et al., 2008; Hanks et al., 2011). However, measuring an 'urgency' signal is non-trivial (Braunlich & Seger, 2016): over the course of a trial, the absolute values of many decision-related factors (such as the level of accumulated evidence) as well as factors which may not be related to the decision (e.g. the ramping of activity in vmPFC) increase. How can we be confident that we have uniquely isolated an 'urgency' signal from the mix? We used

several techniques in an attempt to control for this. In the neural analysis, we included in our regressions a quantity that started high and logarithmically decreased – akin to a neural adaptation signal – as well as the quantity that started small and logarithmically increased, akin to an actual urgency signal. By including the former as an additional regressor, we could ensure that unique effects of general trial ramping were included in the adaptation signal, and by taking the log, we ensured that signals of interest were not just those with a general increase or decrease in activity (as may be expected from the type of spurious signal described). We also note the importance of interpreting this signal alongside behaviour; combining the clear behaviour evidence with the neural analysis adds extra strength to our interpretation of both.

Notably, the drift diffusion model (DDM), which postulates that evidence is accumulated linearly until it reaches a bound, provided a poor fit to the reaction time data in this environment. The DDM has been highly successful in explaining reaction times in humans and non human primates, and it has a wealth of neural support in particular from single unit recordings. Why then does it fare badly as an explanation here? We suggest three causes. The first pertains to the orthogonal influences of evidence strength and evidence variability on trial difficulty in our experiment. Previous studies have often used a measure of "coherence" to determine trial difficulty, whereby the variability itself is the determinant of evidence strength, and thus the two are not dissociable. The explanatory power of the DDM in its current form seems ill-suited to generalise to scenarios such as the one we employed, in which a number of influences can lead to an identical evidence state (decision variable). Furthermore, previous work has often focused on fitting the DDM to reaction time distributions (though see (Kang, Petzschner, Wolpert, & Shadlen, 2017)). Our task also allowed us to interrogate the data with sharper precision: due to the entirely participant-paced nature of the experiment, the decision to view more information, as well as to commit to a decision, were both active choices. This allowed us to have high certainty at any given time the exact statistics of the decision information seen so far and at time of choice, and as such we were able to predict outcomes on a sample to sample basis, and to develop heretofore untested models that we put to the test using these rich data. Finally, it ought also to be noted that the DDM and the adaptive gain model are in nature very similar, with the key difference being the nonlinear weighting of information during integration.

A utility-maximising agent, when faced with the choice of whether to gather more information, or to commit to a choice now, should simulate the value of all possible future states and commit at the

point that will yield the highest expected (Drugowitsch et al., 2012). However, whether or not such demanding computations are plausible for a biological agent remains questionable. We created a simplified version of such a forward-simulating model: a 'greedy' model that calculated the expected value of the sample one step into the future, and committed on the current sample if the opportunity (expected value of taking another sample) was negative. However, despite its normative motivation, this model failed to produce a pattern of performance that was close to human performance.

Across the field, much is known about what decision people will make, and factors – optimal or otherwise – that influence choices. However, decisions are only good when they are made at the appropriate time: assessing every menu option of every café in town to find the perfect lunchtime sandwich will result in me making the best selection, but not until well after the cafes have shut for the day. In order to be worthwhile, choices must be executed in a timely manner, and relatively less is known about the factors that drive decisions to a close. Here we described human behaviour via a mechanism that executes choices when evidence, scaled by its consistency to that seen so far, reaches a bound that collapses over time to avoid prolonged deliberation under conditions of low certainty. By elucidating the mechanism, this opens an avenue to understanding what makes decision timing go "wrong", such as overly-impulsive choices associated with some mental illnesses, and opens the possibility of determining targeted interventions for such problems.

Supplementary material

Alternative models

The SPRT is optimal in the sense that it determines the minimum reaction time required to achieve a given level of accuracy. However the fully optimal solution as to when to respond can only be derived via dynamic programming. Such a process involves simulating a potentially limitless number of future steps in order to calculate the expected value at each one. Although this seemed an unlikely explanation of the mechanisms humans were using, we nevertheless wanted to rule out the idea that a simpler model, that simulated the immediate future and used this to determine whether or not to stop, provided a superior fit to the human data. To this end, we noted that agents can maximise their reward per unit time by taking into account the relative value of responding now and later. The expected value (EV) of responding at time t after information onset can be described as follows:

$$EV_t = \frac{p(cor)_t \times R^+ + q(cor)_t \times R^-}{t + ITI + t_0}$$

Where R^+ and R^- are the rewards and penalties associated with correct and incorrect choices, $p(cor)_t$ and $q(cor)_t$ are the probabilities of correct and incorrect choices at time t, where $q(cor) = 1 - p(cor)$, ITI is the interval interposed between trials, and t_0 is a fixed latency auxiliary to the decision process that is inevitably added to response times.

We considered a predictive model – the one-step lookahead model (OLM) – that estimated the EV of the next sample, and committed to choice when the immediate opportunity cost became negative:

$$Opportunity = EV_{t+dt} - EV_t$$

Unlike the SPRT model considered above, the OLM had full knowledge of the condition-wise structure of the task. Starting with a flat prior across conditions, on each sample the model used Bayes rule to estimate the probability of the sample being drawn from each of the 18 conditions (3 levels of mean

and 3 levels of variance = 9 conditions x 2 categories [CW/ACW]), marginalised to find p(CW), and updated the prior at each step. The model then used these estimates to calculate the expected value of the current sample. In order to calculate opportunity on sample k, an estimate of the evidence that will be provided by sample k+1 is required. To this end, the model used all values of X from 1 to k to estimate the mean and variance of the distribution from which that trial's stimuli were being drawn (unconstrained by the knowledge of the condition-wise structure). It then derived a specific estimate of the next sample by drawing a sample at random from this estimated distribution. As with sample k, the probability of making a correct response on sample k+1 was calculated using Bayes rule (incorporating the prior estimated from sample k). The OLM committed to a choice at the first point that the opportunity of further sampling became negative.

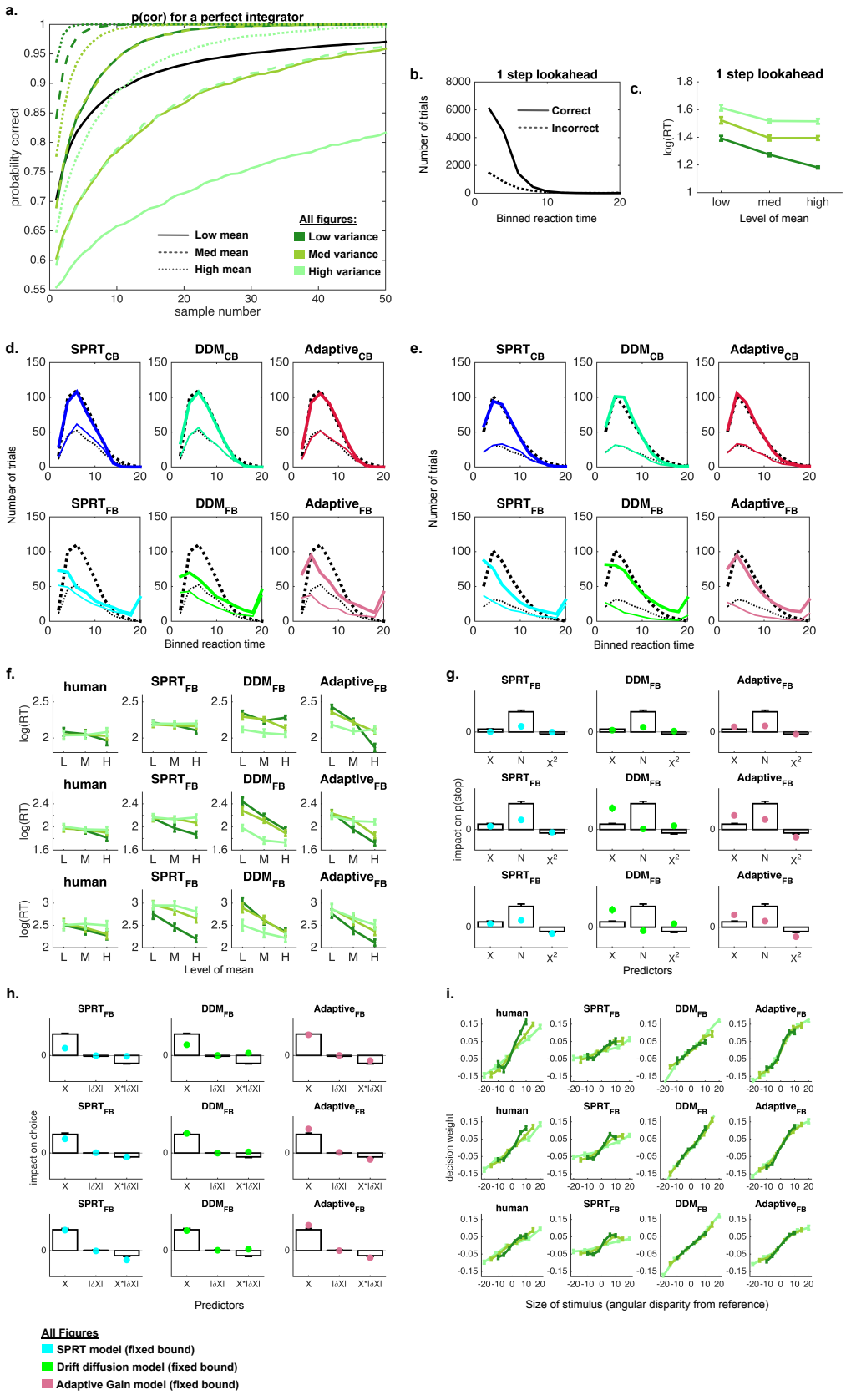
a.
p(cor) for a perfect integrator
probability correct
sample number
Low mean
Med mean
High mean
All figures:
Low variance
Med variance
High variance
b.
1 step lookahead
Number of trials
Correct
Incorrect
Binned reaction time
c.
1 step lookahead
log(RT)
low
med
high
Level of mean
d.
SPRT_CB
DDM_CB
Adaptive_CB
SPRT_FB
DDM_FB
Adaptive_FB
Number of trials
Binned reaction time
e.
SPRT_CB
DDM_CB
Adaptive_CB
SPRT_FB
DDM_FB
Adaptive_FB
Number of trials
Binned reaction time
f.
human
SPRT_FB
DDM_FB
Adaptive_FB
log(RT)
L M H
Level of mean
g.
SPRT_FB
DDM_FB
Adaptive_FB
impact on p(stop)
X N X²
Predictors
h.
SPRT_FB
DDM_FB
Adaptive_FB
impact on choice
X |δX| X*|δX|
Predictors
i.
human
SPRT_FB
DDM_FB
Adaptive_FB
decision weight
Size of stimulus (angular disparity from reference)
All Figures
SPRT model (fixed bound)
Drift diffusion model (fixed bound)
Adaptive Gain model (fixed bound)

Figure 4.S. A. Probability correct as a function of time elapsed for a perfect integrator who chooses on the basis of the sign of the running decision variable. Solid/dotted lines reflect different levels of evidence strength (mean), colours reflect levels of evidence variance, illustrating the dissociable influence of these two factors. **B.** Model distributions of stopping times for correct (solid black line) and error (dotted black) trials, for a one step lookahead model. **C.** Average (log) stopping times (in samples) for trials with low, medium and high mean, i.e. distance to the reference (x-axis) and low, medium and high variance, i.e. distance to the reference (dark, medium and pale green respectively) for the one step lookahead model. **D** and **E.** Human distributions of stopping times for correct (thicker black line) and error (thinner black) trials, and the corresponding fits of fixed- and collapsing-bound variants of the DDM, SPRT and adaptive weighting models (thick and thin coloured lines) for experiments 1 and 2 respectively. **F.** Subjective versus objective weighting of the decision evidence for the fixed bound models, separately for the three conditions of variance (low, medium and high in pale, medium and dark green respectively). The impact of more extreme angles, and angles from more variable streams, is lower in the human data. **G.** Coefficients for statistics predicting next-sample commitment based on sum(X), samples elapsed (N) and sum(X^2) for the fixed bound models. White bars are humans; each panel shows the fit of a model (blue, green and red dots reflect SPRT, DDM and AG models respectively). Rows 1-3 are experiments 1-3 respectively. **H.** The impact on p(choose clockwise) of sample information (X), the prediction error ($|\delta X|$), and the interaction of the latter two for the fixed bound models. **I**. Subjective versus objective weighting of the decision evidence, separately for the three conditions of variance (low, medium and high in pale, medium and dark green respectively) for the fixed bound models.

5. General Discussion

5.1 Summary

The processes underlying human evidence accumulation and decision termination are complex and multi-faceted. Here, we have built on previous work to continue the long journey towards understanding the behavioural and neural mechanisms underpinning perceptual decision making. Across a series of three studies, we have seen that human observers prioritise certain types of information when collecting evidence to inform a decision, and that they terminate choices when the sum of this selectively-integrated information reaches a threshold that collapses over time. This allows a limited-capacity observer to focus resources on diagnostic pieces of information, and to curtail this integration process either when the information is sufficiently supportive of one option over others, or (via the collapsing bound) when evidence is so weak that the benefit of continued deliberation is likely to be negative.

The urn-and-balls task used in chapters two and three, in which participants were required on each trial to indicate from which of two urns a series of eight samples of evidence (variously-sized handfuls of coloured balls) were drawn, could be solved in at least three ways. The two urns that participants had to choose between comprised mainly category A information, or mainly category B information respectively. Thus firstly, participants could take into account the *ratio* of evidence, A:B, in each sample (handful of coloured balls), irrespective of the size of the sample from which the evidence was drawn. This would mean a sample that contained an information ratio of (for example) 2:5 in favour of category A:B would have an identical impact on choice irrespective of the size of the sample (number of balls drawn) that gave rise to this ratio. Secondly, participants could integrate information with respect to its *reliability*, taking into account the evidence in proportion to the information it conveys. A sample that is larger than another (i.e. a sample that contains more balls overall) conveys more information, even if the ratio of evidence is the same in both the larger and the smaller sample. The optimal (binomial) solution to the task accounts for the reliability of information, explicitly factoring the sample size *n* into its calculation of a decision update. Thirdly, rather than taking into account the evidence in proportion to the information it conveys, individuals could apply some weighting such that some information is integrated with higher gain. This is what we term *selective prioritisation* (or *integration*) here.

In chapter two, using a simple version of this task, in which each category (A or B) was characterised by just one 'feature' (colour, in this case), I described evidence counter to strategy one (*ratio-based* information integration) and in favour of strategy two (*reliability-based* information integration): human performance was best captured by a model that summed the evidence on each sample in proportion to the information it conveyed, via a simple mechanism that emulated the binomial solution. In other words, in this experiment, the objective weight of evidence was consistently linearly related to the subjective weight that participants gave the information, as we would expect from a reliability-based solution.

In chapter three, I described a paradigm that made the urn-and-balls task slightly more complex by characterising each category with *more than one* feature, or 'subcategory' (again, each feature/'subcategory' was a different colour). This time, performance was best captured by a model that followed the third strategy, *selectively integrating* information about the more prevalent feature (subcategory) within each of the categories. (In chapter 2 in contrast, the reliability-based solution was equivalent to adding up the number of dots over samples, with no selectivity). In chapter four, using a different type of categorisation task, I detailed a model that again weighted certain pieces of information more strongly than others based on their consistency with the running tally of evidence. Thus together, we never saw a ratio-based strategy, we saw reliability-based (consistent) weighting in one task, prioritising of more prevalent information in another, and prioritising of information consistent with the running mean in the third.

5.2 Linking the results in a common framework

The prioritising of certain pieces of information seen in the latter two experiments, unlike the reliability-based solution seen in the first, can lead to inconsistent decisions. Pieces of information are not treated in a manner that is simply proportional to the information they convey: an identical piece of evidence will be treated very differently depending on the information with which it is presented. The fact that we have seen evidence both for consistency and inconsistency during evidence integration leads to an obvious question: why do we see selective prioritisation of information in some circumstances, but not others? This is not the only query to arise from this pattern of results. The

nature of the gain allocation in the latter two experiments described was quite different: one suggested that humans prioritised the more extreme information (the more prevalent information on a given sample), whilst the other suggested humans prioritised "inlying" information that was consistent with the running mean of evidence. Thus a second question is: can we explain these differences in the nature of evidence prioritisation within just one framework?

Despite these questions suggesting that the evidence across the experiments is potentially disparate, I will argue in what follows that the 'inconsistencies' are actually consistent, linked by a common aim: to allocate gain to the most informative evidence on a given task, subject to capacity constraints.

Why do we see reliability-based behaviour only in some circumstances?

Focusing first on the urn-and-balls tasks, and starting with the first question – why do we see reliability based behaviour in some circumstances, and asymmetric prioritisation in others? – one potential explanation pertains to the fact that the single-feature categorisation task seen in chapter two, and the multi-feature categorisation task used in chapter three differed in the level of cognitive load they produced, which presumably would increase as the number of features per category increased. Thus given that processing capacity may have been less limited by the simpler task used in chapter two, full (reliability-based) integration of the information was possible. However, when cognitive demand increased in the multi-feature task, selective prioritisation of more prevalent information occurred due to the limited capacity of the system under this higher requirement. Similarly, the task employed in chapter four – orientation categorisation of a series of samples drawn from distributions of angles with different levels of mean and variance – was cognitively demanding in the sense that sequentially occurring information from within the same trial could be highly disparate: when the underlying distribution had high variance, one sample could strongly favour category A, and the next strongly favour category B, and so on. Thus some trials featured conflicting information, the integration of which may lead to higher computational demand, and hence again selectivity of gain allocation to inlying evidence was required for the system to cope. Thus the presence of asymmetric weighting in some circumstances and absence in the others is consistent with the notion that selectivity is driven by limited processing resources. Under conditions of high late noise – noise that arises from sources internal to the observer, rather than in the stimulus itself, for example due to this high cognitive demand – capacity constraints mean that not all information can be fully integrated, and therefore

the best decisions can be made by allocating resources to the information that is most diagnostic of a choice.

How can the different types of selectivity be explained?

The second question is an interesting one: why should selectivity under late noise lead to prioritising of inliers in some circumstances (chapter 4, and the previous work described there) and the prioritising of extreme information in others (chapter 3, and the previous work described there)? The answer to this, it has been argued, pertains to the *efficiency* of resource allocation: resources are most usefully deployed when preferentially distributed to process information that is more likely to occur. This efficient resource allocation can be likened to the strategic placement of fielders during a game of rounders in an average school P.E. class: the most likely place the ball will go when bowled is straight past the batter, so it is most effective to have the person best at catching stationed as backstop, with several other fielders placed near base one ready to catch the ball and get the batter out. The likelihood of the ball going anywhere much beyond the bases, on the other hand, is slim, so only one person need be allocated as a 'deep fielder'. In chapter four, the stimuli were drawn from *Gaussian* distributions; therefore participants' prioritisation of inlying rather than outlying information can be explained via this efficiency: allocating resources to extreme stimuli (those far from the distribution mean, and that are therefore unlikely to occur) is wasteful, whereas focusing resources on the most likely stimuli (close to the distribution mean) makes best use of the available resources.

Now consider the scenario in which the evidence is being drawn from a *uniform* distribution: all values within the given range are equally likely to occur. In this case there is no a priori reason to prioritise resources to any one feature value over another, and thus no way to allocate resources in the same way. To continue the playground game analogy, in a game of Bulldog, it is equally likely that the line of players will take any route within the boundaries past the bulldog to the other side. However, when the whistle goes the bulldog can still *efficiently* allocate their attention by focusing on the general area of space to which more of the players head first, neglecting to cover places where fewer people are running, as it is this that will make the bulldog more likely to catch someone. In chapter three, the number of balls drawn from the urn was uniform, and people prioritised information 'on the fly' – based on the most prevalent information on each sample – that allowed them to focus on the most informative evidence in this way. Thus again, the common aim outlined above– to allocate gain to the

most informative evidence on a given task, subject to capacity constraints – is satisfied within these different findings, as all results suggests that resources are being focused in the most efficient way (Li et al., 2017; Spitzer, Blankenburg, & Summerfield, 2016).

However this leaves questions open for further research. Via what mechanism do people come to determine the nature of the underlying distribution of evidence and use it to allocate gain effectively? It is not implausible that they are able to make such deductions very quickly: research into summary statistics has shown that people are able to classify the average feature value (for example, size) of a group of shapes more quickly than it would be possible to process each shape individually, suggesting that humans may perform an initial summary 'sweep' of the scene that picks up on the statistics of the information present (Ariely, 2001; Chong & Treisman, 2003, 2005; Marchant & de Fockert, 2009; Solomon, Morgan, & Chubb, 2011). Furthermore when in doubt, the law of large numbers means that if a feature is sampled often enough, the feature values will form a Gaussian distribution, thus this is a heuristic people could use to infer the underlying structure of commonly occurring stimuli.

Two different types of tasks were used in this thesis: the urn and balls task is what I term a *comparison* task, in which all information for both categories is present on screen at the same time in a given sample, whilst chapter four featured a *categorisation* task, in which only one piece of information is present on the screen at a time, and the information it conveys for category A is 1 minus the information it conveys for category B. The former task type always featured stimuli drawn from a uniform distribution, whilst the latter featured stimuli drawn from a normal distribution. In terms of the allocation of limited resources, it is unclear what would happen if a *comparison* task that featured stimuli from a *normal* distribution was used: we would arguably expect enhanced processing of information that is close to the mean of its respective category, but it is unclear whether the cognitive demand of such a process would be prohibitively high. Thus another potential explanation for the difference in resource allocation seen in the two task types is simply linked to processing capacity: selective integration occurs on categorisation tasks not because of the distribution of the stimuli, but because the cognitive load is too high for an alternative gain control mechanism. It could be that humans have a hierarchy of processes deployed according to cognitive load: reliability-based integration occurs when there is limited load, prioritisation of gain towards expected information occurs when the load is higher, and selective prioritisation of the most prevalent information occurs when the load is higher still. It happens to be the case that the comparison tasks only featured stimuli

drawn from uniform distributions, and the only categorisation task featured stimuli drawn from Gaussian distributions, making it impossible at present to tease apart the two explanations. Thus it remains to be clarified empirically which of the two explanations best accounts for the deployment of processing resources. Of course, it could be the case that neither explanation is superior: other models not yet considered in my analysis may best capture the data.

5.3 Interpreting new results

Urgency

The main focus thus far has been on the mechanisms underpinning evidence accumulation, given this was a theme common to all three experiments. In chapter four, the gain control mechanism that preferentially integrated evidence that was more consistent with the running mean formed part of a wider framework that linked the level of this accumulated quantity to choice times in a self-paced task. In this framework, decisions were triggered when the cumulate of this information reached a threshold level, which differed between individuals, and which decreased over time to allow decisions to be triggered even when evidence remained weak if enough time has passed. This type of signal is often referred to as an 'urgency' signal (Hawkins et al., 2015).

Much previous work that has investigated what drives decisions to a close when the choice is self-paced has focused on fitting model parameters to the distribution of reaction times for correct and error trials. Whilst this process formed part of our model-fitting procedure, we also used a design that enabled us to know precisely at each stage of an ongoing decision (trial) the underlying statistics (for example, the level of cumulative evidence at each timepoint). Our design also required participants to request each piece of information, rather than seeing it automatically, giving us a high degree of confidence that we had a precise measure of whether a person wished to continue sampling or commit to a choice, as both were active processes. This richness of information meant that we were able to interrogate the data in detail beyond model fits to reaction time distributions, by using a general linear model approach to predict the likelihood of decision commitment at any given time based on these quantities. This approach consistently demonstrated that time elapsed in itself was central in causing decisions to be terminated, a finding reflected in the fact that models with a collapsing bound captured data better on every measure used. We hope that the nature of the paradigm used here helps to bolster the case for the importance of an urgency signal in closing difficult

decisions, which we believe is central to understanding what drives commitment in such decisions, but which remains a controversial topic within the literature (Hawkins et al., 2015)

Importantly, building on work using single unit recording (e.g. (Churchland et al., 2008)) we also identified a potential neural substrate of an urgency signal: the lateralised difference in beta band activity – a signal that has been linked to preparation of a motor response (Donner et al., 2009) – was significantly influenced by the amount of time elapsed in a decision. Interestingly, the cumulative decision information had no significant influence on this lateralised activity. This is puzzling given that, whilst the models suggest that time elapsed is important in driving choices, it is clearly not the only factor: these models assume that the level of decision information is what triggers the threshold, as time elapsed alone is not what terminates choices. One option for future research is to repeat this experiment but each participant, after completing the self-paced trials, is shown the exact same trials in a shuffled order under a non-self-paced paradigm. If the signal seen in the beta band is truly an urgency signal that is causally linked to the choice to terminate a decision, we should see it in the self-paced but not the yoked experimenter-paced version of the task. This would help us to understand further the nature of this lateralised signal, and why it may or may not contain decision information.

Subcategory selectivity

An outstanding question from these results pertains to consistencies with the extant literature, rather than the questions above which focused on discrepancies within the experiments themselves: in previous work related to that seen in chapter three, in comparison tasks, information about the more prevalent *category* was preferentially integrated. Here, for the first time, a multi-feature categorisation task was used and the results suggested that in this paradigm, prevalent *subcategory* information was processed with enhanced gain. How this result stands with respect to the current literature warrants some discussion. The use of a multi-feature task to investigate selective integration was a new approach: previously, binary decisions between categories which themselves were comprised of only one feature (such as bar length) have been used. Here instead the decisions required people to choose between categories that were themselves comprised of more than one feature (colour): are the balls drawn from an urn that is mainly comprised of category A, itself comprised of orange and green balls, or mainly category B, itself comprised of pink and blue balls? Using this new paradigm, the results suggested that the more prevalent *feature* (or 'subcategory') within each category – e.g. orange or green – rather than information from the overall category (e.g.

category A, green and orange), was the element that was selectively integrated on the basis of its prevalence. Whilst we may have expected to see information prioritisation at the category level too, this was the first time that this new paradigm, which was designed to explore subcategory selectivity, was used, and was in many ways an exploratory investigation. Arguably the more surprising aspect of this pattern of integration was that, unlike selective integration at the category level, this process in itself did not protect decisions against the influence of late noise. This is particularly interesting given that the beneficial effects of selective integration in the face of late noise has been a key reason for understanding why individuals may behave in line with the process; it is more challenging to account for a mechanism that occurs with no apparent benefit relative to general down-weighting of all information in order to conserve neural resources. However, given that this is the first attempt to investigate selective integration in a multi-feature task, the robustness of this effect ought to be established with further empirical investigation before trying to draw strong conclusions about what this result may mean for the selective integration process in general.

5.4 Future directions and wider questions

An important focus for future work should arguably be to refine further the definition of 'late noise'. Drugowitsch and colleagues made significant progress in formally identifying this source of noise that has a dissociable impact on choice from that of 'early' (sensory) noise. Furthermore, they were able to distinguish two different sources of late noise – integration noise and selection noise – which they showed are increased by increasing the number of items to be integrated or the number of choice options respectively (Drugowitsch et al., 2016). However, increasing the number of options or samples is not the only way to increase the demands of these processes: integration could also be made more challenging by, for example, increasing the disparity in the pieces of evidence to integrate (as in the high variance trials in chapter four), and choices could be made more challenging by, for example, requiring different options to be selected with different modalities (say, a keyboard press for one option and a verbal report for another). However, whilst the notion that increased integration and choice demands will increase late noise at these stages is intuitive – the extent of 'late noise' will often naturally be linked to task difficulty in cognitive tasks –there is not one 'common scale' of late noise, as it can be increased or decreased in multiple ways. This presents a challenge to future research which seeks to link the level of late noise and the nature of selective integration processes seen. One way of approaching this problem consistently is to conduct a set of experiments in which cognitive

load is systematically manipulated in a number of different ways, and to assess the level of late noise generated by these manipulations in an identical way across all tasks so that the level of late noise can be estimated in advance in future experiments.

The finding that, under certain distributions of stimuli, people down-weight perceptually surprising information (such as outliers) is itself in some ways surprising: phenomena such as repetition suppression, expectation suppression (Todorovic & de Lange, 2012) and the oddball response (Picton, 1992) may all suggest that we would do otherwise. All these phenomena are linked by a tendency to allocate more neural resources to 'surprising' stimuli: the oddball response is the enhanced neural response to rare (target) stimuli over frequent (non-target) stimuli. Repetition suppression is the dampened neural response to stimuli which have been seen recently, and expectation suppression is the attenuation of responses that are predicted (and therefore unsurprising). In chapter three, more prevalent – and therefore, statistically less 'surprising' – information is processed with *enhanced* gain, in contrast to these three phenomena. Why do we see these conflicting patterns of results? The traditional distinction between "top-down" and "bottom-up" attention is crucial here: top-down attention is the internal process that guides attention based on prior knowledge and current goals, whereas bottom-up attention refers to external factors that grab attention, such as the salience of a stimulus relative to its context (Katsuki & Constantinidis, 2014). In the urn and balls paradigm in chapter three, no one piece of information would have stood out as inherently salient in comparison to the rest of the information, and as such should not have generated any bottom-up response. In contrast, the task goal was to determine which was the more underlying urn, and therefore top-down attention would arguably have been responsible for the larger neural response to more prevalent – i.e. unsurprising, but highly task relevant – information. Relatedly, in chapter four, the task involved a deliberate manipulation of the variance of the streams from which the stimuli were drawn, meaning that stimuli could be highly different from the running mean and as such may contain an element of bottom up as well as top down salience. This was reflected in the EEG response, consistent with the previous work that shows bottom-up salience leading to larger neural signals: we identified a unique portion of the EEG signal that was attributable to prediction error (i.e. how perceptually surprising the information was): the larger the discrepancy between the current stimulus and the running mean, the higher this neural response.

5.5 Optimal decisions?

The notion of what constitutes a 'good' choice was discussed extensively in the introduction. A number of apparently 'suboptimal' behaviours outlined there can be accounted for within the frameworks provided by these tasks. Particularly, a recurring 'suboptimal' behaviour described in the literature is that of preference inconsistency: a truly rational observer should always show the same preference towards information of a certain value. In both chapters three and four, I showed that selective prioritisation of information goes some way to accounting for this inconsistency: the value or impact of a piece of information on a choice depends on the information with which it is presented. I defined a 'good' decision as one that leads to high accuracy in the face of late noise whilst accounting for the fact that we have limited processing resources. If we allow, based on the arguments presented throughout, that selective integration leads to 'good' choices in this way, then we can conclude that this apparent suboptimality seen in other experiments is simply a by-product of a generally beneficial, resource-efficient process.

This definition of a 'good' decision comprises two subcomponents: resource efficiency, and robustness in the face of late noise. One potentially interesting question is whether the late noise protection arise as a by-product of a resource efficient process, or whether efficiency itself is a useful consequence of robust choices? Selective integration of information does not seem to be an automatic process, as it is not seen in the work presented in chapter two, which could suggest that it is driven by higher task demands (and thus supports the idea that limited resources drive the behaviour). On the other hand, irrespective of whether it saves processing resources, selectivity has been shown to lead to better choices than non-selectivity (though not at the subcategory level): indeed, the fact that selectivity was seen in chapter three despite not conferring any extra benefit to decisions could be taken as evidence that efficiency is the by-product of a semi-default process. Ultimately, this could perhaps be a 'chicken-and-egg' style question and either way, it is the combination of efficiency and robustness, rather than each in isolation, that underpins our definition of a 'good' choice.

Whether or not we accept this way of defining 'good' decisions, our motivation ought not to be an attempt to 'prove' whether or not humans make optimal choices, but instead to understand, given that particular behaviours occur, *why* it happens (Rahnev & Denison, 2018). If decisions that appear

'bad' are made, what is the reason for this, and particularly, is it plausible that we have evolved for millions of years to make fundamental errors in quotidian choices? Of course, this is possible, but a key focus of my work is that 'bad' choices may be i) a by-product of what is generally a good procedure ii) due to our limited capacity system being physically incapable of making 'optimal' choices by some definitions or iii) some combination. All three of the experiments were linked by efficient policies that, generally, lead to accurate choices, despite not always being in line with an 'optimal' observer.

The opening of this thesis states that "good things come to those who weight". Are we now satisfied that this is true? We have seen across a series of three experiments that humans weight evidence according to a variety of features, and that this general policy leads to 'good' choices in terms of maximising accuracy of difficult decisions. This selectivity can be understood within a more general 'gain control' framework, in which individuals prioritise the most informative evidence, based on the statistics of the environment. In the other sense of the word, individuals must wait an appropriate amount of time before choosing – continuing to sample evidence for too long is costly, not sampling for long enough leads to bad choices. We saw that individuals make their decisions when weighted evidence reaches a threshold that collapses over time; the weighting means that choices are more likely to be accurate, the collapsing bound stops individuals from deliberating forever. Thus we can reasonably conclude that good things come to those who…

References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nat Neurosci, 7*(10), 1057-1058.

Aitchison, L., & Lengyel, M. (2017). With or without you: predictive coding and Bayesian inference in the brain. *Curr Opin Neurobiol, 46*, 219-227.

Akrami, A., Kopec, C. D., Diamond, M. E., & Brody, C. D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature, 554*(7692), 368-372.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol, 14*(3), 257-262.

Alais, D., Morrone, C., & Burr, D. (2006). Separate attentional resources for vision and audition. *Proc Biol Sci, 273*(1592), 1339-1345.

Ariely, D. (2001). Seeing sets: representation by statistical properties. *Psychol Sci, 12*(2), 157-162.

Ariely, D. (2009). *Predictably Irrational, Revised and Expanded Edition: The Hidden Forces That Shape Our Decisions*. New York: Harper Collins.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent Arbitrariness": Stable Demand Curves Without Stable Preferences. *The Quarterly Journal of Economics, 118*(1), 73-106.

Baccus, S. A., & Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron, 36*(5), 909-919.

Bartlett, N. R. (1965). Dark and light adaptation. In C. H. Graham (Ed.), *Vision and visual perception. New York: John Wiley and Sons, Inc.; 1965.* New York: John Wiley and Sons, Inc.

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature, 456*(7219), 245-249.

Boehm, U., Hawkins, G. E., Brown, S., van Rijn, H., & Wagenmakers, E. J. (2016). Of monkeys and men: Impatience in perceptual decision-making. *Psychon Bull Rev, 23*(3), 738-749.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn Sci, 11*(3), 118-125.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev, 113*(4), 700-765.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.

Bowman, N. E., Kording, K. P., & Gottfried, J. A. (2012). Temporal integration of olfactory perceptual evidence in human orbitofrontal cortex. *Neuron, 75*(5), 916-927.

Braunlich, K., & Seger, C. A. (2016). Categorical evidence, confidence, and urgency during probabilistic categorization. *Neuroimage, 125*, 941-952.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: WIley.

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science, 340*(6128), 95-98.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol Rev, 100*(3), 432-459.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nat Rev Neurosci, 13*(1), 51-62.

Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., de Gardelle, V., Herce-Castañón, S., et al. (2014). Adaptive gain control during human perceptual choice. *Neuron, 81*, 1429–1441.

Chevalier, G., & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci, 13*(7), 277-280.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Res, 43*(4), 393-404.

Chong, S. C., & Treisman, A. (2005). Statistical processing: computing the average size in perceptual groups. *Vision Res, 45*(7), 891-900.

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nat Neurosci, 11*(6), 693-702.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science, 248*(4962), 1556-1559.

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput Biol, 10*(1), e1003441.

Dayan, P., & Solomon, J. A. (2010). Selective Bayes: attentional load and crowding. *Vision Res, 50*(22), 2248-2260.

de Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proc Natl Acad Sci U S A, 108*(32), 13341-13346.

De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., & Love, B. C. (2017). Social Information Is Integrated into Value and Confidence Judgments According to Its Reliability. *J Neurosci, 37*(25), 6066-6074.

Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr Biol, 15*(6), 543-548.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods, 134*(1), 9-21.

Ditterich, J., Mazurek, M. E., & Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nat Neurosci, 6*(8), 891-898.

Doll, B. B., & Frank, M. J. (2009). The basal ganglia in reward and decision making: computational models and empirical studies. In J. C. Dreher & L. Tremblay (Eds.), *Handbook of reward and decision-making*. Oxford: Academic Press.

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J Neurosci, 31*(16), 6188-6198.

Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res, 1299*, 74-94.

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol, 19*(18), 1581-1585.

Doyle, J., O'Connor, D., Reynolds, G., & Bottomley, P. (1999). The robustness of the asymmetrically dominated effect: buying frames, phantom alternative and in store purchases. . *Psychology and Marketing, 16*(3), 225-243.

Doyle, L. M., Yarrow, K., & Brown, P. (2005). Lateralization of event-related beta desynchronization in the EEG during pre-cued reaction time tasks. *Clin Neurophysiol, 116*(8), 1879-1888.

Drugowitsch, J., DeAngelis, G. C., Angelaki, D. E., & Pouget, A. (2015). Tuning the speed-accuracy trade-off to maximize reward rate in multisensory decision-making. *Elife, 4*, e06678.

Drugowitsch, J., DeAngelis, G. C., Klier, E. M., Angelaki, D. E., & Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. *Elife, 3*.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *J Neurosci, 32*(11), 3612-3628.

Drugowitsch, J., Wyart, V., Devauchelle, A. D., & Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429-433.

Fairhall, A. L., Lewen, G. D., Bialek, W., & de Ruyter Van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature, 412*(6849), 787-792.

Faisal, A. A., & Wolpert, D. M. (2009). Near optimal combination of sensory and motor uncertainty in time during a naturalistic perception-action task. *J Neurophysiol, 101*(4), 1901-1912.

Farashahi, S., Rowe, K., Aslami, Z., Lee, D., & Soltani, A. (2017). Feature-based learning improves adaptability without compromising precision. *Nat Commun, 8*(1), 1768.

Fellows, L. K. (2006). Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain, 129*(Pt 4), 944-952.

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn Sci, 17*(3), 134-141.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci, 17*(1), 51-72.

Frank, M. J., Seeberger, L. C., & O'Reilly R, C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science, 306*(5703), 1940-1943.

Frazier, P., & Yu, A. J. (2008). Sequential hypothesis testing under stochastic deadlines. . *Advances in neural information processing systems, 20*, 465-472.

Friston, K. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci, 360*(1456), 815-836.

Furman, M., & Wang, X. J. (2008). Similarity effect and optimal control of multiple-choice decision making. *Neuron, 60*(6), 1153-1168.

Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage, 19*(1), 200-207.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn Sci, 5*(1), 10-16.

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron, 36*(2), 299-308.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu Rev Neurosci, 30*, 535-574.

Gould, I. C., Nobre, A. C., Wyart, V., & Rushworth, M. F. (2012). Effects of decision variables and intraparietal stimulation on sensorimotor oscillatory activity in the human brain. *J Neurosci, 32*(40), 13805-13818.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Griffin, D., & Tversky, A. (1992). The weighting of evidence and the determinants of confidence. *Cognitive Psychology, 24*, 411-435.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychol Sci, 17*(9), 767-773.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci, 10*(1), 14-23.

Hanks, T. D., Ditterich, J., & Shadlen, M. N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat Neurosci, 9*(5), 682-689.

Hanks, T. D., Kiani, R., & Shadlen, M. N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *Elife, 3*.

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J Neurosci, 31*(17), 6339-6352.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *J Neurosci, 28*(10), 2539-2550.

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E. J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *J Neurosci, 35*(6), 2476-2484.

Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature, 431*(7010), 859-862.

Herce Castañon, S., Bang, D., Moran, R., Ding, J., Egner, T., & Summerfield, C. (2018). Human noise blindness drives suboptimal cognitive inference. *Biorxiv*.

Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Front Hum Neurosci, 3*, 31.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends Cogn Sci, 13*(12), 517-523.

Howe, C. Q., & Purves, D. (2002). Range image statistics can explain the anomalous perception of length. *Proc Natl Acad Sci U S A, 99*(20), 13184-13188.

Howe, C. Q., & Purves, D. (2005a). The Muller-Lyer illusion explained by the statistics of image-source relationships. *Proc Natl Acad Sci U S A, 102*(4), 1234-1239.

Howe, C. Q., & Purves, D. (2005b). Natural-scene geometry predicts the perception of angles and line orientation. *Proc Natl Acad Sci U S A, 102*(4), 1228-1233.

Huang, Y., & Rao, R. P. (2011). Predictive Coding. *Wiley Interdisciplinary Reviews: Cognitive Science,, 2*, 580-593.

Huk, A. C., Katz, L. N., & Yates, J. L. (2017). The Role of the Lateral Intraparietal Area in (the Study of) Decision Making. *Annu Rev Neurosci, 40*, 349-372.

Huk, A. C., & Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J Neurosci, 25*(45), 10420-10436.

Hunt, L. T. (2014). What are the neural origins of choice variability? *Trends Cogn Sci, 18*(5), 222-224.

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F., & Behrens, T. E. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nat Neurosci*.

Jones, M., & Love, B. C. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behav Brain Sci, 34*(4), 169-188; disuccsion 188-231.

Josephs, O., Turner, R., & Friston, K. (1997). Event-related f MRI. *Hum Brain Mapp, 5*(4), 243-248.

Juni, M. Z., Gureckis, T. M., & Maloney, L. T. (2016). Information sampling behavior with explicit sampling costs. *Decision (Wash D C), 3*(3), 147-168.

Kang, Y. H. R., Petzschner, F. H., Wolpert, D. M., & Shadlen, M. N. (2017). Piercing of Consciousness as a Threshold-Crossing Operation. *Curr Biol, 27*(15), 2285-2295 e2286.

Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: different processes and overlapping neural systems. *Neuroscientist, 20*(5), 509-521.

Katz, L. N., Yates, J. L., Pillow, J. W., & Huk, A. C. (2016). Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature, 535*(7611), 285-288.

Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *J Neurosci, 33*(50), 19434-19441.

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron, 84*(6), 1329-1342.

Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci, 28*(12), 3017-3029.

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science, 324*(5928), 759-764.

Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci, 2*(2), 176-185.

Kira, S., Yang, T., & Shadlen, M. N. (2015). A neural implementation of Wald's sequential probability ratio test. *Neuron, 85*(4), 861-873.

Kording, K. P. (2007). Decision theory: what "should" the nervous system do? *Science, 318*(5850), 606-610.

Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427*(6971), 244-247.

Kording, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends Cogn Sci, 10*(7), 319-326.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci, 13*(10), 1292-1298.

Lennie, P. (2003). The cost of cortical computation. *Curr Biol, 13*(6), 493-497.

Li, V., Herce Castanon, S., Solomon, J. A., Vandormael, H., & Summerfield, C. (2017). Robust averaging protects decisions from noise in neural computations. *PLoS Comput Biol, 13*(8), e1005723.

Lieder, F., Stephan, K. E., Daunizeau, J., Garrido, M. I., & Friston, K. J. (2013). A neurocomputational model of the mismatch negativity. *PLoS Comput Biol, 9*(11), e1003288.

Liu, T., & Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *J Neurophysiol, 106*(5), 2383-2398.

Lo, C. C., & Wang, X. J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci, 9*(7), 956-963.

Louie, K., & Glimcher, P. W. (2010). Separating value from choice: delay discounting activity in the lateral intraparietal area. *J Neurosci, 30*(16), 5498-5507.

Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: from biological to behavioral decision-making. *Curr Opin Behav Sci, 5*, 91-99.

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proc Natl Acad Sci U S A*.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat Neurosci, 9*(11), 1432-1438.

Ma, W. J., Navalpakkam, V., Beck, J. M., Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nat Neurosci, 14*(6), 783-790.

Malhotra, G., Leslie, D. S., Ludwig, C. J. H., & Bogacz, R. (2017). Overcoming indecision by changing the decision boundary. *J Exp Psychol Gen, 146*(6), 776-805.

Marchant, A. P., & de Fockert, J. W. (2009). Priming by the mean representation of a set. *Q J Exp Psychol (Hove), 62*(10), 1889-1895.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods, 164*(1), 177-190.

Marr, D. (1981). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman and Company.

Maunsell, J. H. (2004). Neuronal representations of cognitive state: reward or attention? *Trends Cogn Sci, 8*(6), 261-265.

Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends Neurosci, 29*(6), 317-322.

Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cereb Cortex, 13*(11), 1257-1269.

Michael, E. (2015). *Dissociable sources of uncertainty in perceptual decision making*.

Michael, E., de Gardelle, V., Nevado-Holgado, A., & Summerfield, C. (2015). Unreliable evidence: 2 sources of uncertainty during perceptual choice. *Cereb Cortex, 25*(4), 937-947.

Moran, R. (2015). Optimal decision making in heterogeneous and biased environments. *Psychon Bull Rev, 22*(1), 38-53.

Mulder, M. J., van Maanen, L., & Forstmann, B. U. (2014). Perceptual decision neurosciences - a model-based review. *Neuroscience, 277*, 872-884.

Murphy, P. R., Boonstra, E., & Nieuwenhuis, S. (2016). Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nat Commun, 7*, 13526.

Naatanen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends Neurosci, 24*(5), 283-288.

Navalpakkam, V., Koch, C., Rangel, A., & Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proc Natl Acad Sci U S A, 107*(11), 5232-5237.

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175-220.

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron, 37*(1), 149-157.

Nienborg, H., & Cumming, B. (2010). Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? *Curr Opin Neurobiol, 20*(3), 376-381.

Niwa, M., & Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *J Neurosci, 28*(17), 4435-4445.

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat Neurosci, 15*(12), 1729-1735.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J Vis, 5*(5), 376-404.

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. *Trends Cogn Sci, 21*(6), 425-433.

Pare, M., & Wurtz, R. H. (2001). Progression in neuronal processing for saccadic eye movements from parietal cortex area lip to superior colliculus. *J Neurophysiol, 85*(6), 2545-2562.

Parrish, A. E., Evans, T. A., & Beran, M. J. (2015). Rhesus macaques (Macaca mulatta) exhibit the decoy effect in a perceptual discrimination task. *Atten Percept Psychophys, 77*(5), 1715-1725.

Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: children's integration of auditory and haptic information is suboptimal. *Dev Sci, 17*(3), 376-387.

Pettibone, J. C., & Wedell, D. H. (2000). Examining Models of Nondominated Decoy Effects across Judgment and Choice. *Organ Behav Hum Decis Process, 81*(2), 300-328.

Pfurtscheller, G., Neuper, C., & Mohl, W. (1994). Event-related desynchronization (ERD) during visual processing. *Int J Psychophysiol, 16*(2-3), 147-153.

Philiastides, M. G., Heekeren, H. R., & Sajda, P. (2014). Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. *J Neurosci, 34*(50), 16877-16889.

Piazza, M., & Izard, V. (2009). How humans count: numerosity and the parietal cortex. *Neuroscientist, 15*(3), 261-273.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*(3), 547-555.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *J Clin Neurophysiol, 9*(4), 456-479.

Pisauro, M. A., Fouragnan, E., Retzler, C., & Philiastides, M. G. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nat Commun, 8*, 15808.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*(6741), 233-238.

Popper, K. (1959). *The Logic of Scientific Discovery*. New York: Basic Books.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annu Rev Neurosci, 13*, 25-42.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat Neurosci, 16*(9), 1170-1178.

Rahnev, D., & Denison, R. N. (2018). Suboptimality in Perceptual Decision Making. *Behav Brain Sci*, 1-107.

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: theory and evidence. *Curr Opin Neurobiol, 22*(6), 970-981.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci, 2*(1), 79-87.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput, 20*(4), 873-922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*, 347-356.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol Rev, 111*(2), 333-367.

Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience, 89*(4), 1009-1023.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu Rev Neurosci, 27*, 611-647.

Rock, I., & Victor, J. (1964). Vision and Touch: An Experimentally Created Conflict between the Two Senses. *Science, 143*(3606), 594-596.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci, 22*(21), 9475-9489.

Rorie, A. E., Gao, J., McClelland, J. L., & Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS One, 5*(2), e9308.

Rudebeck, P. H., Buckley, M. J., Walton, M. E., & Rushworth, M. F. (2006). A role for the macaque anterior cingulate gyrus in social valuation. *Science, 313*(5791), 1310-1312.

Rushworth, M. F., Kennerley, S. W., & Walton, M. E. (2005). Cognitive neuroscience: resolving conflict in and over the medial frontal cortex. *Curr Biol, 15*(2), R54-56.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci, 27*(3), 161-168.

Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology, 32*, 135–168.

Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *J Vis, 11*(12), 13.

Spitzer, B., Blankenburg, F., & Summerfield, C. (2016). Rhythmic gain control during supramodal integration of approximate number. *Neuroimage, 129*, 470-479.

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science, 304*(5678), 1782-1787.

Summerfield, C., Behrens, T. E., & Koechlin, E. (2011). Perceptual classification in a rapidly changing environment. *Neuron, 71*(4), 725-736.

Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron, 59*(2), 336-347.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci, 11*(9), 1004-1006.

Summerfield, C., & Tsetsos, K. (2012). Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms. *Front Neurosci, 6*, 70.

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends Cogn Sci, 19*(1), 27-34.

Sun, J., & Perona, P. (1998). Where is the sun? *Nat Neurosci, 1*(3), 183-184.

Thura, D., Beauregard-Racine, J., Fradet, C. W., & Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. *J Neurophysiol, 108*(11), 2912-2930.

Thura, D., & Cisek, P. (2014). Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making. *Neuron, 81*(6), 1401-1416.

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J Neurosci, 32*(39), 13389-13395.

Trueblood, J., & Pettibone, J. C. (2015). The phantom decoy effect in perceptual decision making. *Journal of Behavioural Decision Making, 30*(2), 157-167.

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proc Natl Acad Sci U S A, 109*(24), 9659-9664.

Tsetsos, K., Chater, N., & Usher, M. (2015). Examining the mechanisms underlying contextual preference reversal: Comment on Trueblood, Brown, and Heathcote (2014). *Psychol Rev, 122*(4), 838-847.

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proc Natl Acad Sci U S A, 113*(11), 3102-3107.

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychol Rev, 117*(4), 1275-1293.

Tsetsos, K., Usher, M., & McClelland, J. L. (2011). Testing multi-alternative decision models with non-stationary evidence. *Front Neurosci, 5*, 63.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*, 105-110.

Tversky, A., & Kahneman, D. (1974a). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1974b). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *The Americal Economic Review, 80*, 204-217.

Urai, A. E., & Pfeffer, T. (2014). An action-independent signature of perceptual choice in the human brain. *J Neurosci, 34*(15), 5081-5082.

van Beers, R. J., Sittig, A. C., & Gon, J. J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *J Neurophysiol, 81*(3), 1355-1364.

Vickers, D. (1979). *Decision Processes in Visual Perception*: Academic Press.

Voskuilen, C., Ratcliff, R., & Smith, P. L. (2016). Comparing fixed and collapsing boundary versions of the diffusion model. *J Math Psychol, 73*, 59-79.

Wald, A., & Wolfowitz, J. (1949). Bayes Solutions of Sequential Decision Problems. *Proc Natl Acad Sci U S A, 35*(2), 99-102.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*, 273–281.

Welch, R., & Warren, D. (1986). Intersensory interactions. In K. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 25.21-25.36): Wiley.

Wyart, V., de Gardelle, V., Scholl, J., & Summerfield, C. (2012). Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron, 76*(4), 847-858.

Wyart, V., Myers, N. E., & Summerfield, C. (2015). Neural mechanisms of human perceptual choice under focused and divided attention. *J Neurosci, 35*(8), 3485-3498.

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A, 109*(9), 3593-3598.

Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature, 447*(7148), 1075-1080.

Yu, A., & Dayan, P. (2004). Inference, attention, and decision in a Bayesian neural architecture. *Adv. Neural Inf. Process Syst., 16*, 1577–1584.

Zylberberg, A., Fetsch, C. R., & Shadlen, M. N. (2016). The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *Elife, 5*.